

When not to Pull Punches

ARE the people of some races inherently less intelligent than others? In the past few years, this contentious question has made all kinds of trouble. In countries such as South Africa, of course, the question is asked without much restraint, and the supposedly self-evident answer is used as a justification of illiberal social policies, at least by a minority of the caucasian population. So it is not surprising that the mere asking of the question has come to be associated with illiberality, and there are many professional scientists who hold that the subject is so full of danger that, for the time being at least, investigations of the inheritance of various attributes by the people of different races should be postponed (see *Nature*, **228**, 1013; 1970 and **232**, 2; 1971). This argument has now been countered by an eloquent plea from a distinguished group of scientists, mostly psychologists and geneticists, including the two conspicuously controversial figures in the field—Professor Arthur Jensen and Professor Hans Eysenck—but with support from others with no particular axes to grind such as Professor Jacques Monod and Professor F. H. C. Crick. The “Resolution on Scientific Freedom Regarding Human Behaviour and Heredity” is published in the *American Psychologist* for July.

Important issues stem from what the signatories of the round robin have to say. First, they complain that those who now emphasise the role of heredity in human behaviour are being subjected to “suppression, censure, punishment and defamation” which is in practice similar to, and in principle the same as, the suppression of Galileo in Italy, the criticism of Darwin in England a century ago, the persecution of Einstein and his works in Nazi Germany and the suppression of Mendelian genetics in Stalin’s Russia two decades ago. They go on to complain that published positions are misrepresented, that attacks on research intended to throw light on the inherited basis of racial differences come from scientists who are committed to the environmentalist view of how observed differences arise as well as from “political militants on campus”, not always scientists. The most serious charge is that there are many scientists who are persuaded of the “great role of heredity in human behaviour” but who prefer to hold their tongues in public for fear of giving offence. The result, they say, is that “a kind of orthodox environmentalism dominates the liberal academy” so that teachers are inhibited from turning to explanations based on inheritance which, in their view, are already valuable and which call for further and more energetic research.

Few will quarrel with the principle that science must be open. To do so is to deny the lessons of recent centuries that the consequences of suppression are invariably confusion and malevolence. Equally, however, there is no doubt that in fields in which the results of research are liable to misinterpretation, the scientific community has a duty not merely to gain further understanding of what are often complicated problems but also to ensure that the results of their research are accurately and moderately interpreted to the world at

large. In other words, there are many circumstances in which research must be seen not merely to be accurate and perceptive but seemly and modest as well. Unhappily, many of those who have argued for the hereditary basis of human behaviour, and especially of intelligence, have been the worst offenders against these simple rules. Dr William Shockley of the California Institute of Technology (not one of the signatories of the resolution now published) has for example done a great deal of harm with his recurring proposals that the National Academy in the United States should investigate the matter and then produce an authoritative statement of the position. For the truth is that only the most dogmatic adherents of the hereditary view would think that the time is ripe for a final assessment of the relative importance of nature and nurture—the old-fashioned terms—in, say, the determination of intelligence. This is why some exponents of the hereditary view—Professor Eysenck, for example, may be held to have been less artful than their own high intelligence would have made possible, in advocating the genetic case. By suggesting that the issue is settled, they have sharpened the conflict between themselves and others. And by doing so, they have helped to stimulate—for that is what the world is like—the way in which more moderate associates (Professor Jensen among them) have been pilloried.

So what should now be done? The first thing to acknowledge is that certain obvious differences between the races do exist. All negroes, for example, have darker skins than all caucasians, at least if the albinos are left out of the count. Anthropologists have long wrestled with the problems of how to categorize other races, and to distinguish between them on phenotypic grounds, and have compiled a great deal of anecdotal detail as well as making a little progress. If the course of natural selection had not been interrupted by the peculiar gregariousness of all the races, and if it had not in any case been comparatively slow because of the long interval between conception and puberty which characterizes all human races, there is of course a possibility that inter-racial differences might have become the foundation for the emergence of distinct species.

Another difficulty is that the racial differences which are of most concern to sociologists are in no sense as simple as those of skin colour between negroes and caucasians. All negroes may have darker skins than all caucasians, but there is no equivalent statement to be made about differences of intelligence or even physical fitness. In the past few years, Professors Eysenck and Jensen have been prominent among those who say that the average intelligence of negroes is less than that of at least some other races, but even if the potential uncertainties in this conclusion are overlooked, the plain fact remains that many negroes turn out to be more intelligent than many caucasians. In the statistical comparison of racial attributes, in other words, the statistical distributions overlap. And in reality, of course, the nurture school is right to argue that when more is known of the subtlety of the ways in which environment can



help to determine intangible characteristics such as intelligence, the extent of the overlap may be much greater than may now be apparent.

These simple truths have a cruel meaning for those who would attempt to use a better understanding of racial differences as a basis for making social decisions. Enlightened societies have learned, sometimes painfully, that it is exceedingly unwise to base social policy on the assumption that some people are more equal than others in any way at all. In circumstances like these, it would be entirely proper that an enlightened government persuaded of the reality of some of the differences of which Professors Eysenck and Jensen have spoken in recent years would be impelled not to illiberal policies intended to accentuate the differences but in the opposite direction, possibly to the deliberate encouragement of miscegenation.

The moral for the signatories of the Jensen resolution is plain enough. They happen to be working in a field of science in which it is not permissible to think that scientists have done their duty when they have described the results of their work to their professional colleagues, and widely (and accurately) popularized them. It is also necessary, in these special circumstances, earnestly to ensure that the audience is clearly aware that the results are usually irrelevant to the management of society as it is at present. The dangers of misunderstanding are so many that a proper interpretation of the facts cannot possibly be left to chance. With this important reservation, unfortunately lacking from the declaration in the *American Psychologist*, the plea for openness is irresistible. Intellectual intolerance is in its way as evil as racial intolerance. But some of the signatories must mend their ways.

Turning-point for BA?

ONE way or another, the annual meeting of the British Association at Leicester next week will determine whether this ancient and recently fuddy-duddy organization will be able to make a continuing claim on professional and public attention. For several years, the association has been struggling to avoid collapse through bankruptcy, and in the process has been too eager to compromise on issues of principle—the annual hand-out of £10,000 from what used to be the Ministry of Technology, now running out, was always likely to disappoint those who thought that the association could exert an important influence in persuading young men and women to turn their attention from pure science to technology. The chief consequence of this and other subventions was to postpone the energetic pruning of the association's office overheads, mercifully undertaken early in 1971 by the body known as the Committee of Review. With all these anxieties, however, the association has paid too little attention to the continuing redefinition of objectives which changing circumstances require. Only now, in the nick of time, has the association put forward a set of proposals for reorganization which promise to provide a framework within which it might hope indefinitely to prosper, intellectually as well as financially.

The General Committee of the association will be asked next week to approve a number of far-reaching proposals, one of which will have the effect of turning

the committee into a kind of electoral college, with responsibility for electing a much smaller council than the elephantine body which is at present responsible for making policy, but which is at once too large and too honorific to be effective. Although the electoral system now proposed is complicated by the need somehow to satisfy the sixteen sections of the association that they will be fairly represented on the new council, the fact that ordinary members of the association would, on the new proposals, have a say in choosing council members will by itself go a long way to ensure that the association does what the members want. The notion that the chairman of the council should be a paid officer of the association, the intellectual and executive head of the organization, is another step in the right direction. And for the next few years at least, there is a prospect that the government will indirectly provide a substantial subvention by means of a grant from the Royal Society.

What should the association try to do with the new lease on life that is in prospect? The annual meeting is certain to continue, for this appears to be what members want. The association will, however, have to work out ways of making the meeting more effective both as a vehicle for the clubby professional talk which does at least serve to bring active researchers and other scientifically inclined people together and as a means of conveying a vision of science and of its implications to the world at large. But an annual meeting is not enough to keep the association together, which is why the plans for carrying out special studies of important matters of public interest, and for holding less elaborate meetings on special topics, are widely to be welcomed. The need for some means by which matters of public concern, from the economics of the crusade against pollution of the environment to the realities of recent progress in biomedical science, should be studied carefully by a body which has the skill and independence necessary to make such work convincing. In short, the British Association is as much needed now as when it was created, nearly 140 years ago. The hope is that the meeting this week at Leicester will appreciate the importance as well as the magnitude of the challenge.

100 Years Ago



We are glad to be able to record the addition to the Brighton Aquarium, of a specimen of Muller's Topknot (*Rhombus hirtus*.) It was netted off the Brighton coast last week. But one capture of this rare fish off the Sussex coast is recorded by Yarrell, but it is more frequently taken off the Cornish coast. The interesting event is also announced of the birth of a young cuttle-fish, which signalised its entrance into the world by an immediate discharge of the sepia fluid.

From Nature, 6, 378, September 5, 1872.

OLD WORLD

Support for the European Research System

A DUAL method of supporting research within universities is essential if research is to flourish according to a recent report published by the Organization for Economic Cooperation and Development (*The Research System, Vol 1, France, Germany, United Kingdom*).

In a detailed inquiry into how research is organized and financed in these three European countries, which each spend approximately the same percentage of their gross national product on research, the OECD concludes that financial support for the institution encourages the university to have a policy towards research whereas financial support for the individual "encourages the value of research".

This endorsement of the European system of providing money for research, which directly contrasts with the system in the United States where individual grants are usually the sole means of support, is further strengthened by the statement that such a system must be assessed on a regular basis which in Britain, as in Germany, is done every five years.

The report also poses pertinent questions about the link between university teachers and researchers. It says that "the model of a teacher-researcher can no longer be defined in the same terms as even fifteen years ago". It points out that research with its need for teamwork and the need to man equipment continually does not leave the scientist much time for teaching. Therefore, according to the report, there is a need to define "a more precise distinction between advanced postgraduate teaching and basic teaching, and between research proper and the administration of a research team".

The report, however, draws a line at separating universities into those that carry out research and those which are entirely devoted to teaching.

The university systems in France, Germany and Britain do not come out completely unscathed and each system is accused of resisting change. In particular the systems are accused of being "sluggish" because they are reluctant to conduct multidisciplinary research.

The method of examining research projects in Britain is complimented as being the "most open". The French CNRS, the report says, makes use of "the least manageable form of committee system" whereas the Germans have "succeeded in establishing committees by research projects".

Science carried out within a government framework has, according to the

report, done all that has been asked of it, but such an arrangement is susceptible to certain failings. There is a risk that the political and administrative atmosphere might conflict with the requirements of fundamental research. Such an atmosphere will turn scientists into civil servants says the report—a transition, incidentally, which in Britain Lord Rothschild and the government in its white paper has been trying to encourage. Even giving the government laboratories and agencies legal autonomy has not been enough to avoid this danger, says the OECD and research within government agencies still

does not have the genuine autonomy it needs.

Even though the report was largely compiled before the upheaval caused by the Rothschild report in Britain some of its recommendations are particularly apt. It points out that in a transition period difficulties are bound to arise for all those who are exposed "to the backlash of experimentation and of trial and error". "In institutions whose structures and whose very purposes are being challenged at all levels, unrest generating discontent among scientists may jeopardize the equilibrium and soundness of research."

ATOMIC ENERGY

Inflation Bites Home at the IAEA

THE International Atomic Energy Agency is running into financial difficulties. In spite of an overall expenditure of \$13 million in 1971, and an approved budget for 1972 of \$15.3 million, certain areas of the agency's work have suffered. Its technical co-operation activities accounted for \$2.12 million of the agency's regular budget in 1972 compared to \$1.25 million in 1970, but the agency declares that the real resources available, once inflation and fluctuating exchange rates have been taken into account, are substantially less. As a result the agency had to turn down 44 requests from member states for assistance this year; in 1971 it only turned down 27 such requests.

Nevertheless the agency had a very active year, executing nine large projects for the United Nations Development Programme, compared with three in 1971. This increase, the agency says, "is to some extent a reflection of the growing ability of nuclear energy centres in developing countries to make a direct contribution to industrial and economic growth". These projects include the construction of a national nuclear energy centre in Chile, exploration for uranium in Greece and Pakistan, and the development of agricultural production through the application of nuclear technology in Brazil.

The agency's training programme is also feeling the pinch, however. This year the number of training fellowships financed through the agency has fallen to its lowest level since 1967—only 288 were awarded, against 384 last year. This sharp drop is due in part to the curtailment of UNDP funds for regional and inter-regional projects; as a result all but one of the 1972 training

courses are being financed from the agency's regular programme and not with UNDP assistance.

The agency also reports that 52 nuclear power reactors were ordered in 1971, with a total capacity of 46,572 MW. The agency says that previous forecasts that nuclear power will account for about 14 per cent of the world's electricity generation in 1980, and about 50 per cent by the turn of the century, still seem valid; by the year 2000 about 3.5 million MW of nuclear generating capacity should be installed.

The agency's planned programme of work for the next five years has also just been released. The bulk of it will be a continuation along familiar lines with expansion where additional finance will permit. More work is expected on radiation attenuated vaccines, and on food that has been disinfected by irradiation, for example.

The agency also sees a substantial increase in the number of developing countries ordering nuclear power plant in the next five years and much more of the agency's time and money will be spent facilitating this. Further, an increased effort will be made to discover and evaluate uranium sources in developing countries, but the intensification in these activities will lead to a cut in spending on reactor technology. Work on advanced nuclear projects, such as ship propulsion and magneto-hydrodynamics, will, however, continue at the present level.

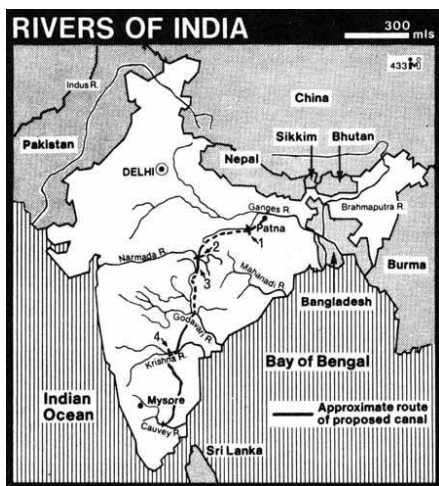
The agency has also published its 1973 proposed budget which it puts at \$18.1 million. Unhappily, the agency fears that most of the increase will again be swallowed up by inflation and changing rates of exchange.

INDIAN SCIENCE

Watering the Plains

By the year 2000, India plans to have built, from scratch, a national water grid that will ferry fresh water 1,650 miles from the shores of the Ganges to Upper Anicut on the banks of the River Cauvey in the depths of Madras. This extremely ambitious scheme, which will involve the construction of 661 miles of canal to hump the water over India's central highlands, springs from the country's unequal rainfall distribution. During the summer monsoons, water pours down the giant rivers of the north and east, flowing away to the sea, while southern, central and western India struggle through the year on markedly low rainfalls. India's answer is to build a north-south canal which will link her major rivers, all of which flow almost due east to the Bay of Bengal or due west to the Indian Ocean.

Thirty years from now, when the scheme will hopefully have been completed, ten million acres of land that is now prone to chronic drought will be under irrigation, drinking water supplies will have been stabilized throughout the country, and low cost canal transportation will link northern India to the south.



1. Sone barrage ; 2. Bargi reservoir ; 3. Tunnel ; 4. Srisaillam reservoir.

The scheme was first mooted last century by Sir Henry Cotton, a British irrigation engineer with an eye on the low costs of canal transport, but its imminent construction is due chiefly to Dr K. L. Rao, minister for Irrigation and Power, who has seen the scheme through a United Nations feasibility study which pronounced it technically possible, and who is now awaiting the results of a UN economic study.

Estimates of the cost vary. The Indian government at present is settling for the lower end of the spectrum, suggesting that the scheme will cost

Rs 22,400 million, but figures as high as Rs 40,000 million have been put forward.

India's annual run off has been calculated at 1,672,600 million cubic metres—about 2.5 times the run off from the Mississippi-Missouri basin—but 85 per cent of India's rain falls in the north in the summer. When the scheme is completed, the flood waters that fill the Ganges to bursting each summer will be tapped at Patna (see figure) and pumped at the rate of 60,000 cubic feet per second to the existing Sone barrage. Fifty thousand cubic feet per second will then be lifted along the Sone river for 150 days each year to the Bargi reservoir. Twenty thousand cubic feet per second will be released into the Narmanda river for the drought affected rivers in its basin, and the bulk of the remainder will flow to the Inchampalli reservoir. From there, 50,000 cubic feet per second will flow under the force of gravity into the Srisaillam reservoir from which it will enter the

Pennar river and flow down to the Upper Anicut on the Cauvery.

In addition to the north-south canal, the construction of which will require driving a tunnel through the mountains south of the Bargi reservoir, a network of secondary canals will take water from the central one to help irrigate a yet wider area of the country.

Power needed to pump the water is estimated at 8.5 million kilowatts, but the natural slope of the rivers will be used to generate 2.3 million kilowatts of this total. Pumping will chiefly be limited to the 150 days during the monsoons when the flood waters in the Himalayan rivers will be used to generate cheap secondary power for pumping.

The scheme is not yet a certainty. If the second UN study shows it to be uneconomic then a major reconsideration of the project will follow, but if all goes well India may, by about the year 2000, have solved its chief irrigation problems.

The End is in Sight

FISCHER played the Pirc Defence in the 17th game of the World Chess Championship, a defence he has not been known to use before in a championship game. This demonstrates once again the thorough preparations that Fischer must have undertaken in the months preceding the match; in contrast, Spassky gives the impression of having taken this phase of the match rather too lightly in his pre-match training.

Spassky dealt adequately with the problems of Fischer's play, however, and at the cost of a pawn managed to work up a dangerous attacking formation. Fischer was obliged to give up the exchange in order to fight for a draw, a result he achieved surprisingly easily.

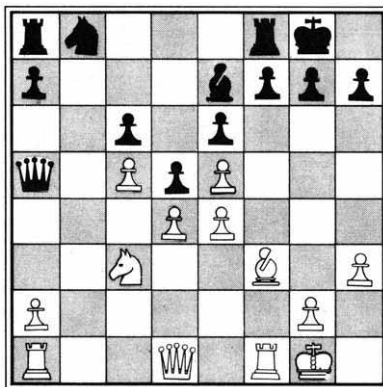
The 18th game was a fiercely fought Sicilian with both players again striving

for a win. A difficult game arose with queens and rooks on the board; Spassky was obliged to give up a pawn, but managed to pin down one of Fischer's rooks to the defence of his king in the corner. The position was fairly evenly balanced when the players repeated moves for a draw immediately after the adjournment.

The 19th game found Spassky on the white side of an Alekhine's Defence, but playing a variation against it superior to the one he had used earlier in the match. The game became rather lively when at the 18th move (see figure) Spassky made a speculative knight sacrifice. The play for the next few moves was highly tactical but Fischer found the right resources to fend off any danger. Indeed, Spassky was perhaps a little fortunate to be able to find a simplification to reach yet another drawn ending.

The play over the last six games or so has been evenly matched following Spassky's improved form, but Fischer, leading 11-8, is now only a stone's throw from winning the match.—J.P.

FISCHER BLACK



NEW WORLD

A Sliver in the Election Platforms

by our Washington Correspondent

Now that the Republicans have gone through the formality of nominating Richard Nixon and Spiro Agnew to run for a second term in the White House and the Democrats have finished the painful process of putting their ticket back together, the serious business of choosing the next President and Congress of the United States is getting under way in earnest. A small sliver, scarcely large enough to call a plank, in the platforms adopted amid the fratricidal soul-searching of the Democratic convention and the slick television performance of the Republican gathering, is how each party intends to manage the scientific and technological resources of the United States.

Science and technology will clearly take a back seat in the coming election, not only because there are more potent issues such as the Vietnam war, the economy and law and order to be discussed, but also because there are no large differences between the science policies of the parties. The rhetoric is the same—both platforms call for science to be used for the public good in the form of new products, improved health care and services in general, and each points out the need to forge new links between government and industry. When science and technology do figure in the election campaign, the debate is therefore likely to centre around whether or not the Nixon Administration has backed its rhetoric with dollars. And another issue is likely to be unemployment among scientists and engineers.

The Democratic platform, for example, accuses the Nixon Administration of allowing the overall level of scientific research in the United States to decline. "For years, the United States was the world's undisputed leader in science and technology", the platform document points out, but "now that leadership is being challenged, in part because of the success of efforts in other countries, and in part because of the Nixon Administration's neglect of our basic human and material resources in this field".

The Republican document, written a month after the Democratic platform, countered, however, that "we have increased Federal efforts in civilian research and development by 65 per cent—from \$3,300 million to \$5,500 million". The problem is that both statements are correct, but they do not tell the full tale. For one thing, although the Nixon Administration has raised the total ex-

penditure on research and development, inflation has more than eaten up the increases, and the purchasing power of the science budget will be no greater in 1973 than it was in 1969 (see accompanying article).

As for the figure of 65 per cent, it is an optimistic estimate based on obligations, not on expenditures, and takes no account of inflation. And so debates about the total level of support for science and technology, which will undoubtedly arise during the election, are likely to sink into semantics, although the Administration's claims have the advantage of simplicity, while the Democrats' attempts at refuting them involve complex arguments.

But the Democrats face an even more difficult task in assailing the thrust of the Nixon Administration's science policy, which has been geared towards focusing science and technology on domestic problems, such as medical research and pollution control. Such a policy is, of course, difficult to fault from a liberal standpoint, and the chief line of attack will probably be that the Administration has not moved fast enough.

The Democratic platform document, for example, suggests that a Democratic Administration would be "eager to take the management methods and techniques devised for the defense and space program, as well as our technical resources, and apply them to the city, the environment, education, energy, transportation, health care and other urgent domestic needs", all of which sounds like the Presidential Message on Science and Technology which was delivered to Congress in March.

But it also sounds like the preamble to Senator Edward M. Kennedy's National Science Policy and Priorities Act, which was passed by the Senate last week and which was cosponsored by Senator McGovern (see *Nature*, 238, 427; 1972).

Apart from seeking to involve the National Science Foundation in all kinds of applied research projects, Kennedy's bill would set up programmes to help retrain unemployed scientists and engineers, and that is one aspect on which the Democratic platform also takes the Nixon Administration to task. "As Democrats", the document claims, "we understand the enormous investment made by the nation in educating and training hundreds and thousands of highly skilled Americans in science and technology. Many of these people are now unemployed, as aerospace and defense

programs are slowly cut back . . . so far, however, the Nixon Administration has paid scant attention to these problems". The Republican platform, on the other hand, says "we pledge expanded efforts to aid unemployed scientists and engineers. We are determined to see that such on-going efforts as the technology mobilization and re-employment program are effective". Once again, there is no difference in sentiment between the two parties, and the issue is likely to centre around the level of commitment.

As for space research, the Democratic platform is notably silent, while the Republican document claims that impressive domestic benefits are already being reaped from the space programme and it also pledges support for the space shuttle. McGovern, along with many other leading Congressional Democrats, has expressed his personal opposition to the shuttle, however, a fact which will not win him many friends among the aerospace workers who are already hard pressed by cut-backs in the space budget, and one which is perhaps best left out of the platform.

Although the two party platforms have similar policy thrusts as far as science and technology are concerned, however, it seems that so far scientists themselves are organizing chiefly in support of Senator McGovern's candidacy. A Scientists for McGovern Committee has been formed, with Herbert York of the University of California, San Diego, and Harry Palevsky of Brookhaven National Laboratory, as cochairmen and Vigdor Teplitz of MIT as secretary.

The committee has recently sent out a mailing to 100,000 scientists and engineers, the response to which, according to Teplitz, has been "satisfactory but not overwhelming". A science and technology policy committee has also been drafting a policy paper for McGovern, which should be ready later this month.

The thrust of the opposition to the Nixon Administration's record on science and technology, Teplitz said last week, is that, although the Administration "knows what the needs are, and has drawn up lists of what needs to be done, they just don't do them". For example, he criticized the Administration for not developing an energy policy, and suggested that the Administration is at fault in concentrating on the fast breeder reactor while neglecting other areas of research such

as nuclear fusion and solar energy.

Finally, Teplitz suggested that "our impression is that science and technology just isn't very important to the Nixon Administration". The Administration would, of course, dispute that claim, but the fact remains that science and technology *per se* is probably not important enough to the electorate to figure in the forefront of the election campaign.

SCIENCE FUNDING

Almost Recovered

by our Washington Correspondent

ONCE again, the National Science Foundation has come up with a set of figures to show that federal expenditures on science and technology are increasing. Total government expenditures are expected to increase from \$15,500 million in 1971 to "an all-time high" of \$17,800 million in 1973, the report suggests. Not so prominently displayed, however, is the fact that if inflation is taken into account, total expenditure in 1972 was some \$2,500 million less than in 1967, and it will still be below the 1967 figure, in purchasing power, in 1973.

Nevertheless, the decline in funding that set in during the late 1960s seems at last to have reached its nadir, even though the increases of the past two years have been insufficient to wipe out the decreases of the previous four. And there has also been a marked shift in funding with the Department of Defense, NASA and the Atomic Energy Commission accounting for a diminishing share of the total research and development budget. In 1963, for example, these three agencies together carried off 90 per cent of the total federal science budget, but in 1973, they are expected to take a mere 75 per cent.

The NSF figures also show that the development end of the research and development spectrum is undergoing a relative decline, while basic and applied science funding are on the upswing. Between 1963 and 1973, the share of the federal science budget devoted to development has dropped from 68 per cent to 59 per cent, while basic and applied research together have increased their share from 32 to 41 per cent. Among the various fields of science, engineering accounts for the largest expenditure, while expenditures on life science have grown the fastest during the past decade.

MENTAL HEALTH

Ecopsychiatry

by our Washington Correspondent

RELATIONSHIPS between pollution and health have for long been the subject of debate and study. The effect of air pollution on respiratory diseases, the contribution of poor sanitation to the spread of diseases such as cholera and even the relationship between soft water and arteriosclerosis have all been discussed in the scientific and popular Press. But what of the effects of pollution on mental health? What, for example, are the consequences to the mind of living in urban squalor, breathing polluted air and eating chemically contaminated food? According to a study carried out for the National Institute of Mental Health, such questions have received at best only scant attention, and much more research is required not only on the physiological effects of pollutants on the central nervous system, but also on the mental stresses and strains of living in a degraded environment.

The study (*Pollution: Its Impact on Mental Health*, DHEW publication No. (HMS) 72-9135, US Government Printing Office, \$0.45), which was essentially a literature search, turned up 110 scientific papers dealing with pollution and mental health. Not surprisingly, most of the research so far has been concerned with physical damage to the central nervous system resulting from exposure to specific pollutants such as mercury, lead and arsenic. More subtle effects, such as the impact on mental health of chronic exposure to noise, poor housing and environmental degradation in general, have so far received scant attention. The report suggests, moreover that even those studies that have been carried out on the direct physiological effects of pollutants often leave much to be desired—for example, the impact on the nervous system of chronic exposure to low levels of insecticides, metals and even airborne lead, is poorly known.

That there are yawning gaps in the literature concerning pollution and mental health is perhaps not surprising in view of the complexities of the interrelationships, and the relative novelty of concern over the environment but, according to Dr Rene Dubos, who wrote an introduction to the report "The elusiveness of the problem is no excuse for ignoring its importance or neglecting its study".

As for the importance of studying the impact of pollution on mental health, Dr Dubos suggests that the "most deplorable aspect of existence in American cities may not be murder, rape and robbery, but the constant exposure of children to pollutants, noise, ugliness and garbage in the street. This constant exposure conditions children to

accept public squalor as the normal state of affairs and thereby handicaps them mentally at the beginning of their lives". But whether or not ecopsychiatry will ever be sufficiently rigorous to contribute to the solution of such problems is a moot point, and fortunately, one which lay outside the purview of the study.

LUNAR SCIENCE

Swansong for Apollo

by our Washington Correspondent

APOLLO 17, the last in the series of flights which fulfilled President Kennedy's wish to have men land on the Moon in the 1960s, is set to blast off from Cape Kennedy at 9.53 pm local time on December 6 (2.53 am December 7 GMT). The preliminary timetable for the mission, released last week by NASA, will make Apollo 17 the longest of all the flights in the series, and up to 21 hours have been set aside for scientific work on the lunar surface. Splashdown is set for 2.24 pm (7.24 pm GMT) on December 19.

The landing site, Taurus-Littrow, is set in the Taurus Mountains in the south-eastern rim of *Mare Serenitatis*. The site is believed to be a young volcanic area, characterized by a smooth appearance and a lack of large blocks, set between massif units of the Taurus Mountains. It is hoped that the landing site will enable both younger volcanic materials and older massif materials to be collected. Work on the lunar surface is scheduled to take place in three seven hour shifts starting at 11.33 pm (GMT) on December 11, 10.13 pm (GMT) on December 12 and 9.33 pm (GMT) on December 13.

Because Apollo 17 is the last in the series, it seemed until recently that there would be no manned space flights from the United States between Skylab, which is set for launch next year, and the advent of the shuttle in about 1978. But the joint agreement with the USSR for a rendezvous between an Apollo spacecraft and a Soyuz spacecraft in 1975 has bridged the gap, and perhaps kept the NASA manned spacecraft teams together. NASA announced last week, for example, that Dr Rocco A. Petrone, Apollo Program Director, has been appointed Program Director for the NASA portion of the joint US/USSR space mission, and that he will use the Apollo team for the mission.

Support staff at the Manned Spaceflight Center in Houston, the Marshall Spaceflight Center in Alabama and the Kennedy Space Center in Florida, have, however, not fared so well. NASA announced last week that to keep its spending within 1973 budget limitations, some of this staff would have to be laid off. Up to 2,700 jobs may be involved.

NEWS AND VIEWS

Phyllosilicates and the Solar System

INTERPLANETARY space contains a considerable amount of solid material, and the Earth on its journey around the Sun continually collides with this debris. Recent observations show that about 1000 ton is being accreted annually. This debris has a mean geocentric velocity of $130,000 \text{ km h}^{-1}$ and so most of it is vaporized by frictional heating in the Earth's upper atmosphere. A few objects are sufficiently large to penetrate the atmosphere without being completely destroyed, and they arrive at the surface of the Earth as meteorites. One of these will hit an area the size of Britain every 2 to 3 years, that is, about 500 reach some part of the Earth's surface each year.

Meteorites contain the most primitive material known. They have remained almost unchanged since the formation of the solar system, and are a thousand million years older than the oldest known rocks on Earth. Meteorites have not suffered the high degree of differentiation that has taken place in the Earth's crust, and thus they are much more representative of the composition of planetesimals, the primitive rock fragments which coalesced to form the planets Mercury, Venus, Earth and Mars and the rocky cores of the outer planets. Asteroids are perfect examples of planetesimals—small irregular rocky bodies similar in form to Phobos, the moon of Mars which was photographed recently by Mariner 9.

Several thousands of these tiny minor planets move in orbits between those of Mars and Jupiter. The total mass of all the asteroids is of the order of 0.03 per cent of that of the Earth, so, unlike the planetesimals which coalesced gravitationally to form the Earth, the mass of the asteroids was too small and they still remain as individual bodies. Meteorites, being asteroidal planetesimals on orbits that intersect that of the Earth, provide the chief clue to the composition of planetary interiors and to the composition of the non-volatile elements in the condensing nebula which formed the planets in the solar system. The study of meteorites forges a vital link with the preplanetary solar system.

By far the most common class of meteorites are the chondrites, which make up more than 75 per cent of known meteorite falls. These stony bodies usually contain chondrules, spherical grains of iron magnesium silicate a few millimetres across, but the most primitive members of this class, the type I carbonaceous chondrites, do not even contain these.

Most cosmochemists have assumed that these type I carbonaceous chondrites are in fact the primitive essence of the solar system and that the abundance of condensable elements in this class of chondrites is the same as the primitive abundance in the original contracting nebula. But certain of the minerals in type I carbonaceous chondrites were not there originally; for example olivine and pyroxene are thought to be mechanical additions and certain sulphates and carbonates are obviously late deposits. Phyllosilicates, on the other hand, have been there all the time.

Phyllosilicates are the most common material in carbonaceous meteorites; they constitute about 65 per cent by weight and occur as submicroscopic foils or flakes, which are typically about $1000 \text{ \AA}$ in diameter, $100 \text{ \AA}$ thick and

polycrystalline. John Kerridge, of the Institute of Geophysics and Department of Chemistry of the University of California, has concluded—in a communication on page 44 of this issue of *Nature*—that it is the phyllosilicates in type I carbonaceous chondrites that are the “primitive essence” of the solar system, and not the whole chondritic material. One of the fundamental problems that must be solved before it can be definitely established that it is only the phyllosilicates that are the basic building blocks of the planets is the inter-relation between these phyllosilicates and magnetite, which makes up about 15 per cent by weight of the meteorite.

Kerridge finds that the phyllosilicate-magnetite association is not the result of a hydration of the original condensates in the nebula in equilibrium with solar gases at temperatures of less than 315 K (that is they were not formed at the same time), but that phyllosilicate is a basic unaltered material which has been mixed up with the altered magnetite actually in the meteorite, and that these two substances have never been in equilibrium.

The reasons propounded are that the phyllosilicate has retained its iron and has an unusually high nickel content, and that it has not been aqueously altered or hydrated since formation; it also contains rare gases of a planetary and not solar wind type, like the magnetite. In other words the silicate and magnetite obtained their gases in different environments and were subsequently mixed at such a low temperature that equilibration did not occur. Also the composition of phyllosilicate matches the abundance of the principal elements in the Sun. A problem arises here, however, because data on solar abundances, particularly in the cases of iron and silicon, are rather uncertain.

Kerridge concludes that the Sun, the planets and the chondritic meteorites were all formed from the same nebular cloud, and that the phyllosilicate material found in type I carbonaceous chondrites is the “primitive non-volatile essence” of the solar system, which formed the rocky cores of all the planets and satellites. In the light of this hypothesis he also concludes that the cosmic ratio of iron to silicon is 0.53 ± 0.06 , a finding that should be of great help to solar astrophysicists in their interpretation of solar spectroscopic data.—D. W. H.

Starting off DNA with RNA

ONE of the most pressing problems of DNA replication has been how the synthesis of new strands of DNA is initiated, for all of the known DNA polymerase enzymes of both bacteria and the cells of higher organisms appear to be able to extend a DNA chain which has already been started; but none can commence synthesis of a chain anew. Although hints have been appearing in the literature for the past few months that RNA may be involved in DNA synthesis, the first direct evidence of such an involvement in replication of the DNA of

Escherichia coli has now been provided by Sugino, Hirose, and Okazaki in the July issue of *Proceedings of the US National Academy of Sciences* (69, 1863; 1972).

DNA polymerases are active when given a primer—a short oligonucleotide sequence complementary to the strand which they are about to transcribe and which is extended by the enzyme. The primer added to start replication for *in vitro* experiments has usually been DNA; but the effect of the drug rifampicin on replication implies that RNA may be transcribed to act as a primer for DNA replication too. Rifampicin prevents RNA polymerase from synthesizing RNA on DNA templates; although it does not inhibit replication, Brutlag, Schekman and Kornberg (*Proc. US Nat. Acad. Sci.*, 68, 2826; 1971) found that it prevents conversion of single stranded circles of phage M13 DNA to the duplex form in cells of *E. coli*. Because chloramphenicol, which inhibits protein synthesis, does not have this effect, the implication is that it is the synthesis of RNA itself—and not its intermediary role in the synthesis of proteins—which is demanded for the initiation of replication.

Following up this discovery, Wickner, Brutlag, Schekman and Kornberg found earlier this year (*Proc. US Nat. Acad. Sci.*, 69, 965; 1972) that conversion of M13 single strands to double helical DNA requires RNA synthesis *in vitro*. A radioactive label in a precursor of DNA is transferred to a degradation product of RNA when the product of these experiments is cleaved with alkali; this suggests that the RNA is used as a primer sequence which is then extended covalently by addition of nucleotides to form a DNA strand. But because no such similar mechanism seems to be needed for replication of another single stranded DNA phage, ØX174, it is difficult to conclude that the use of RNA is widespread.

One essential experiment to confirm the need for the mediation of an RNA primer is to show that ribonuclease H—an enzyme which specifically destroys the RNA part of DNA-RNA hybrids—prevents conversion of single strands to duplexes. Otherwise it remains possible that RNA synthesis is implicated in only a subsidiary and not a direct way. A requirement for the presence of RNA has been shown in this way by Keller (*Proc. US Nat. Acad. Sci.*, 69, 1560; 1972), but in a system which relies upon replication of *E. coli* DNA by a polymerase derived from human tumour KB cells or from another bacterium. This may mean that the use of RNA as a primer extends to all organisms; or it may be a peculiarity of this heterologous system.

RNA may be used to initiate new cycles of replication in *E. coli* cells, for Lark (*J. Mol. Biol.*, 64, 47; 1972) has found that replication cannot start when RNA synthesis is inhibited by the addition of rifampicin up to ten minutes before DNA synthesis is due to start. Stopping protein synthesis at this time has no effect, so that although the cells have produced the proteins which they need to start a new cycle of replication, they cannot synthesize DNA if they are prevented from transcribing RNA. The obvious role for this RNA is to act as a primer for DNA synthesis, although it may of course have some other role—especially as the essential transcription step seems to take place ten minutes before DNA replication commences.

But there is one important flaw in the conclusion that DNA synthesis can continue if it is started by transcription of an RNA primer. DNA synthesis takes place by a discontinuous mechanism in which new DNA is synthe-

sized in short fragments—the Okazaki segments—which are only subsequently joined together into a covalent DNA chain. How are these fragments to be initiated? Their synthesis cannot take place by continuation of an enzyme along its template once it has started, for the frequent interruptions in the integrity of the new DNA must ensure that synthesis starts anew with each fragment.

An explanation for this paradox is provided by Okazaki and his colleagues in their latest triumph; each Okazaki segment starts with a primer sequence of RNA. The buoyant density of a radioactive label incorporated into the fragments depends upon its length; by using a very short labelling time, Sugino *et al.* have found that the fragments are in fact heavier than would be expected if they are composed of DNA alone. But when the extracted DNA is treated with reagents such as alkali or ribonuclease, which destroy RNA, the fragments show the characteristic density of single strands of DNA.

Direct evidence for the presence of RNA is provided by experiments which show that a precursor of RNA is incorporated into the fragments, although it is removed very rapidly from them. This implies that the RNA primer is removed from its DNA partner very soon after commencement of synthesis of the DNA attached to it. The noticeable effect of the RNA portion of the Okazaki fragments on their density suggests that RNA must comprise a relatively large length of the very nascent molecule; Sugino *et al.* suggest that the RNA may be 50 to 100 nucleotides long (a completed fragment is of the order of 1000 nucleotides).

The idea that RNA synthesis is continually required for DNA replication raises almost as many problems as it solves. The experiments which pointed to a role for RNA in the initiation of cycles of replication did not suggest that RNA synthesis is needed for the continuation of DNA synthesis; how does initiation of the very first fragment differ from that of the later fragments? Another problem is that the rate of RNA synthesis is many times slower than the rate of DNA synthesis; does transcription of primer sequences take place during replication—in which case it may be a rate limiting step—or are pre-formed sequences used, in which case it is difficult to see how they would be produced?

If transcription takes place during replication, there may be frequent signals along the bacterial chromosome to direct RNA polymerase to initiate and terminate synthesis of the primer sequences; this would mean that many short sequences of DNA must be devoted to control rather than to coding. How would these sequences be interposed amongst the genes themselves? An alternative model is to suppose that the synthesis of the RNA sequences is random in location, perhaps caused by some mechanical link between the replication apparatus and transcription. The failure of rifampicin to inhibit replication itself suggests the intriguing possibility that some enzyme other than RNA polymerase may catalyse the transcription of RNA for replication; but all extractions of *E. coli* cells for RNA synthesizing capacity have produced only this one enzyme.

The RNA starter sequences must be very rapidly removed from the chromosome by nuclease action and the gaps filled. This requires two additional enzymes to participate in replication. Do bacteria contain an enzyme activity analogous to ribonuclease H, which has so far been found only in the cells of higher organisms, or does one of the known ribonucleases have this role?

This may be one of the most difficult questions to resolve, for it has so far proved impossible even to identify the ribonuclease which degrades messenger RNA. And how are the gaps filled? One candidate for this role must be DNA polymerase I, for the *E. coli* mutants lacking this enzyme appear to join their Okazaki fragments together more slowly than do wild type cells. Another enzyme involved may be DNA polymerase II, which is often found in those membrane systems which catalyse replication.

Indeed, the implications for attempts to synthesize DNA *in vitro* may be amongst the most important. The need for RNA synthesis may make it difficult, or even impossible, for purified enzyme systems to replicate DNA *in vitro*; one reason why they have often failed to do so successfully may be that, although their capacity to extend DNA chains is retained, they cannot initiate synthesis of the Okazaki fragments. Okazaki's latest discovery is almost as challenging as his original discovery that the synthesis of DNA is discontinuous.—From our Molecular Genetics Correspondent.

MULTIPLE CROPPING

Research in India

from a Correspondent

THE latest results of an extensive research programme into the whole problem of multiple cropping, which was initiated by the Indian Agricultural Research Institute in 1966, could be of tremendous significance for increasing agricultural production, especially in developing countries with high rural unemployment (*Research Bulletin*, new series, No. 8 (IARI, New Delhi, 1972)).

Effective utilization of solar energy and soil nutrients for crop production is only partially achieved even under the most advanced or intensive systems of farming. Even with a controlled supply of water and adequate solar radiation throughout the year, the life span of single crops from sowing to harvest does not usually allow land to be continuously occupied in successive or rotational cropping. Nevertheless, systems of multiple cropping are being effectively used in countries like Taiwan, China, Japan and Egypt as well as in India.

The principal objective of the Indian programme was to increase the economic yield of crops from each unit of irrigated land, largely by telescoping the production time of several complementary crops in sequence on the same land.

This was made possible in two ways—the development of a relay system of cropping and by important contributions from plant breeding.

The relay system involves interplanting the next crop before the previous crop is harvested. In this way a sequence of, say, mung (a pulse crop), maize, potato (or oil seed *Brassica*) and wheat, which would normally take more than 365 days to mature, can be harvested within the year. By judicious choice of crop and variety a wide range of three and four crop relays have been successfully produced.

As the optimum yield over a short period of time is of paramount importance for multicropping, a different set of criteria for plant breeding is required. With many crops, particularly cereals, the trend is to increase the harvest index, that is weight of harvested product/total plant weight of the crop, and this is of special significance for multicropping in that an over-abundance of vegetative parts tends to delay maturity. A photosynthetically efficient but smaller type of plant would also allow for a higher number of economically productive units per area of land.

Although the chief consideration in the research programme of the Indian Agricultural Research Institute was to increase production per unit area over a shorter time scale, it was appreciated that, for the systems to be successful, minimum soil cultivation and economical use of water are important. A secondary objective was to ensure that the physical, chemical and biological properties of the soil are maintained, and if possible improved. Again, by judicious choice of crops and varieties due regard has been given to the problem of pests and diseases without resorting to protection by chemicals.

Of special significance for India is the fact that, although many of the crop relays provided significantly greater production and net profit a hectare, the intensification also implies a dramatic increase in employment potential. Thus the difference between a fallow wheat rotation and a mung-maize-potato-wheat relay amounts to an increase in production from 5 to 14 t of food grain or equivalent a hectare and an increase in manpower requirement from 85 to 432 man days a hectare a year.

Economic growth is often achieved at the expense of reduced employment of manpower, but the adoption in practice of the methods advocated by the Indian Agricultural Research Institute will, it is claimed, provide employment for an extra 17 million people to alleviate the distressing amount of unemployment in the rural areas of India. The stimulus arising from a well coordinated programme of agricultural research is contributing significantly towards providing India and other developing countries

with increased production of several crops in a way that will help to alleviate both famine and rural unemployment.

SOLID STATE

Bends in TiNi

WHEN a sheet of the alloy TiNi is bent beyond its elastic limit at less than about 160° C (the martensitic transition region) the temperature of the distorted part increases, but when the sheet is straightened again heat is reabsorbed and the temperature drops by as much as 15° C. This is one of several novel properties of the alloy—its memory is already well known—that are described by Wang and Buehler in a recent issue of *Applied Physics Letters* (21, 105; 1972).

Wang and Buehler also point out for the first time a curious effect that takes place when a TiNi wire below its transition temperature is heated locally after it has been stretched just beyond its so-called limit of easy plastic flow. At the point where the heat is applied the wire bends so that one straight part is at an angle of about 30° to the other. What is particularly interesting is that further heating at the same place causes the wire to straighten again. If the heat source is moved along the length of the wire, however, the bend will follow it and can be made to disappear off the end. Wang and Buehler say that these and other features of TiNi support the contention that its martensitic transition is chiefly of an electronic nature.

RIBOSOMES

Roles in Ribosomes

from our Molecular Biology Correspondent

SINCE the heady triumphs of total reassembly, relatively little has occurred in the ribosome field to stir the spirit. There has been instead a long period of consolidation, or painstaking and unspectacular programmes of fractionation, sequencing and characterization, and the signs now are that the pay-off is at last at hand. In particular a start has been made in defining the structural and functional roles of individual proteins of the *E. coli* ribosome. In a brief but action-packed paper, Funatsu and Wittmann (*J. Mol. Biol.*, 68, 547; 1972), for example, define the chemical basis of streptomycin resistance in *E. coli*. Both streptomycin resistance and dependence are associated with a single protein of the 30S subunit, S12. This protein was isolated from nine streptomycin resistant mutants of four allele types, and examined by tryptic fingerprinting. The mutant proteins differed in each case from the wild type

in only one peptide, and between the nine mutants the changed peptide was one of two. In both cases a lysine was replaced; one of the two lysines gave place variously to arginine, asparagine or threonine, the other always to arginine. The substitution is, moreover, characteristic of the allele type, regardless of the strain. From the sequence data so far available, it appears that the two susceptible lysines occur at positions 42 and 87 in the chain, which agrees nicely with the conclusions of genetic mapping experiments. The substitutions in all cases correspond to simple-base mutations. In a streptomycin-dependent mutant, the substitution is again at position 42 where a glutamine residue now appears. The allele types differ in the degree of translational fidelity of which the ribosome is capable and it thus seems that this function depends profoundly on the residue in this position.

The isolation in Wittmann's laboratory of the set of proteins from both the large and small subunits has opened the way to a variety of other approaches to the assignment of specific roles. An example is to be found in the identification by Gray *et al.* (*Eur. J. Biochem.*, **28**, 412; 1972) of the proteins implicated in anchoring the 5S RNA, which is an integral part of the large subunit. Gray and Monier had reported earlier that in the presence of a fraction of the group of proteins that are extracted at defined salt concentrations from the ribosome (the so-called "split proteins"), 5S RNA would associate with 23S to form a complex, retained by membrane filters. Two-dimensional gel electrophoresis now allows an identification of the proteins in the relevant fraction, and by examining the ability of separated proteins of the large subunits, singly and in combination, to cause the retention of a 1:1 complex of 5S and 23S RNA by a membrane filter, the identity of the binding proteins could be established. The upshot is that one protein (L18) is indispensable, but not in itself sufficient for complex formation. When either of two other proteins (L6 and L25) was added, binding ensued. Another protein (L2) also stimulates the retention of the RNA by the filters, but this result has no counterpart in sucrose gradient sedimentation, which suggests that L2 may merely exert some effect on the RNA so as to increase the interaction of the complex with the filter. Of the three essential proteins only one has a primary binding site on the 23S RNA, whereas the other two will bind to 5S RNA. All three proteins are released from the ribosome, together with the 5S RNA, at high salt concentrations. The same proteins appear contrariwise to be necessary for the reincorporation of 5S RNA into the unfolded large subunit, and it

therefore seems not at all unlikely that they are part of the attachment site in the native particle.

Gray and Monier have also shown that part of the 23S RNA can be removed without disturbing the 5S binding site, and they have now made some progress towards the hallowed end of identifying a (presumably multiple) protein binding site on the 5S RNA (*FEBS Lett.*, **24**, 156; 1972). When a complex, containing unlabelled 23S RNA, labelled 5S RNA and protein, is exposed to ribonuclease a considerably smaller proportion of labelled nucleotides is released than from the uncomplexed 5S RNA. The residual hot RNA, when extracted and subjected to nucleotide mapping, then reveals which parts of the chain are protected in the complex. The largest piece to be found extends from residue 69 to the 3' end (residue 120), and contains also the 5' terminal sequence from residues 1 to 11, which is complementary to, and presumably base-paired with, the tract at the 3' end, to form a kind of stem. In most of the fragments, however, this stem is missing, because of a break at residue 110. The rest of the molecule contains a sizable internally complementary tract, with breaks at three

bonds, at the presumed turn. It seems likely then that the important protein binding sites occur in the sequence 69-110, though parts of the chain are still accessible to the nuclease.

A completely different approach to the assignment of functions to particular proteins in the ribosome involves the use of covalent affinity labels. Czernilofsky and Kuechler (*Biochim. Biophys. Acta*, **272**, 667; 1972) have bound phenylalanyl-tRNA, modified at the amino-acid with a reactive and radioactively labelled p-nitrophenyl-carbamyl group. In the presence of poly U, and only then, this binds to the 50S subunit, leaving a strong label in two ribosomal proteins, after digestion with ribonuclease. A definitive identification of the two proteins, which can be presumed to lie in or near the ribosomal A-site, has not yet been achieved. Pellegrini, Oen and Cantor (*Proc. US Nat. Acad. Sci.*, **69**, 837; 1972), using the same strategy, labelled only one protein. Whether this was a consequence of the different specificity of their reactive derivative (bromoacetylphenylalanyl-tRNA) or, as Czernilofsky and Kuechler suggest, the low levels of radioactivity incorporated, is not clear.

Pulsations in White Dwarf Stars

THE application of modern statistical techniques, such as power spectrum analysis, to astronomy has resulted in an explosion of information about variable objects of all kinds. These techniques have been applied by various groups to the study of objects as diverse as the Sun and quasars; now, however, a continuing flow of data is coming from a systematic study of rapidly varying stars being made by one group. In next Monday's *Nature Physical Science* (September 4), B. Warner and E. Robinson report that studies at the McDonald Observatory, Texas, have resulted in the discovery of five "new" periodicities in the cataclysmic variable systems. These short period binaries show a variety of variations in their light curves, including considerable random contributions to the flickering caused by mass transfer. So this discovery of five periodic white dwarfs could only have been possible using a combination of sophisticated statistics and the best optical systems available.

Some idea of the importance of this combination is provided by the figures Warner and Robinson quote for the level at which visual inspection of light curves can detect periodicities (1 per cent of intensity) compared with the capability of power spectrum analysis to pick out periodic variations of only 0.001 mag in an 18 mag object (provided long data runs are available). The

objects in which periodic variations have now been detected are Z Cam, CN Ori, UX Uma, AH Her and AM CVn (HZ29); the latter object is discussed in some detail, by the same authors, in the current issue of the *Monthly Notices of the Royal Astronomical Society*.

These results are important because they lend weight to the theory that the white dwarf component in these binary systems is likely to be oscillating in a non-radial mode, as suggested by Chanmugam (*Nature*, **236**, 83; 1972). Indeed, Warner and Robinson believe that they have identified discrete modes of oscillation, and that their observations show transitions between these modes. It seems likely that such detailed information about the behaviour of white dwarfs in close binary systems will not only reveal details of the structure of white dwarfs themselves but will also provide an insight into the processes causing the outbursts of these cataclysmic variables, which are at present far from completely understood. The classic example of such a system is WZ Sge; it is gratifying that the development of a theory about the evolution of the archetype (see, for example, Faulkner, *Astrophys. J. Lett.*, **170**, L99; 1971) and improved observations of other members of the same family seem to be leading towards the same coherent picture of cataclysmic variables.

MATERIALS

On Boom and Doom

from our Materials Science Correspondent

A RECENT public venture by Harvey Brooks, professor of applied physics at Harvard University, was a lecture on materials in a steady state world, delivered last autumn to the two leading professional bodies in American metallurgy, and now published in their joint journal (*Metallurgical Trans.*, **3**, 759; 1972). Professor Brooks is a man of unusually wide experience in the borderlands of science, engineering and politics in the United States, and has participated in many enquiries on various aspects of science policy.

He began with a thumbnail sketch of Forrester's world models (the lecture antedated publication of *The Limits to Growth*, which took that model a stage further), and contrasted the underlying outlook of that form of global future gazing with the changes during the past century in the supply and real price of raw materials, bearing in mind the fact that dire prophecies of the exhaustion of natural resources have been current for a long time—they were particularly prevalent in the first two decades of this century. Among a number of striking observations, he showed that the consumption of materials *per capita* had grown much less rapidly than the gross national product (GNP) *per capita*, and that the output for each man employed has risen particularly fast in the mineral extraction industries. He made the intriguing suggestion that one symptom of the lack of pressure on these industries has been the sharp decline of mining engineering and extractive metallurgy in universities.

Having very fairly presented the optimist's case, Professor Brooks then turned to the other side of the medal. Some materials pose a much more serious problem than others: thus *per capita* use of paper is closely correlated with *per capita* GNP, but grows five times for each doubling of the *per capita* GNP. Moreover, the price in real terms of forest products has trebled over the past century, whereas on the same basis the prices of minerals are unchanged. Professor Brooks went on to classify metal ores in a particularly illuminating way. Some metals, especially iron and aluminium, are present in enormous quantities in ores whose grade varies from those at present exploited, continuously down to very much lower levels. Others, such as Cu, Ni, V, Pb, occur in ores with a continuous range of grades, including substantial supplies in ores at present somewhat too lean to use, and these will become steadily dearer without, presumably, a state of crisis being reached. The real problem comes with what he calls the "metal vitamins", used

as essential alloying elements; these include, among others, Hg, W, Ta, Ag, Sn, Mn and Mo; some of these of course have "nutrient" as well as "vitamin" uses. They are won from rich mineral veins which are fast being exhausted, but once this has happened, there is apparently no supply of progressively leaner ores, only the general background concentration in the Earth's crust, which is orders of magnitude lower.

Many metals can be won for a long time by a combination of improved technology and higher real price, but this will not be true for all. When one of the problem metals also poses a special pollution problem, it will be interesting to see how rapidly substitution or abolition of function results. Mercury is a good example, and it may be significant that the price has recently been plummeting: on the other hand, the influences of cartels among producers and consumers are so strong that it is dangerous to read too much into price fluctuations of metals.

Professor Brooks continued by outlining some of the more novel problems of the future, which are all connected with the increasing difficulty of discerning trends far enough in advance to plan the necessary industrial changes in the extractive industries. He instances sudden unpredictable political changes, including ever stronger cartels of suppliers, and changes imposed by rapidly veering public opinion, which do not allow the necessary lead time for technical change. He also assessed in general terms the problems arising out of the universal drive towards growth, and said quite bleakly that "there

appears to be a major discontinuity between current trends and national objectives, in virtually all national economies, and the need to obtain some sort of equilibrium with our environment. It is not obvious that the various homeostatic mechanisms of the world economy will operate to set in motion automatically the forces which will propel us towards a new equilibrium. The discontinuity between current trends and goals and ultimate necessities is very hard to think about in a logically consistent way. The apocalyptic rhetoric of the ecological ideologies is not very helpful because they do not tell us how to get there from here, or even how to begin".

In a tentative attempt at propounding some solutions, he urged a government financed expansion of long range geochemical and geophysical research "aimed at defining and assessing the future availability of metallic ores", research which cannot fairly be left to mining firms. He thinks that, although substitution of more common for less common materials must continue apace, whole technologies must give way to less materials-intensive ones: he gave as an example advanced telecommunications as a substitute for travel. He also emphasized the role of draconian changes in national laws to oblige all companies to do what no one firm could do in isolation without going out of business, and instanced the new American car exhaust emission laws as an example of what can be achieved. This law may well in retrospect mark a watershed, showing many who now indulge in fashionable despair what man can do to help himself.

Factor Regulating Translation of RNA Coliphages

IN *Nature New Biology* next week (September 6) Revel and various colleagues describe in two reports the discovery of a new control element which seems to play a part in regulating the translation of particular cistrons of a messenger RNA and which also comprises one of the polypeptide chains of phage Q β RNA replicase. The new factor is designated by Groner, Pollack, Berissi and Revel as factor i because it appears to interfere with the translation of the coat protein cistron of the RNA phage MS2 by binding to the chain initiation factor 3.

In the absence of interference factor i, initiation factor 3 promotes the efficient translation of the MS2 coat protein cistron *in vitro*, but when factor i is added synthesis of coat protein is markedly reduced while translation of the RNA synthetase cistron is enhanced. Factor i also apparently stimulates the translation *in vitro* of some phage T4 cistrons while inhibiting the translation of others so that the net result is little

or no change in the total protein synthesis; there is, however, a change in the relative proportions of particular proteins. Clearly factor i has the properties of an element which might play an important role in regulating the expression of the *Escherichia coli* genome and the genomes of coliphages.

But perhaps even more intriguing than the role of factor i in translation is the discovery, reported by Groner, Scheps, Kolakofsky and Revel, that factor i, a polypeptide with a molecular weight of 74,000 daltons, is one of the three polypeptide chains specified by the host *E. coli* genome which, together with the chain specified by the Q β RNA synthetase cistron, comprise the Q β RNA replicase. Apparently factor i first promotes the translation of the synthetase cistron of an RNA phage and then becomes incorporated into the synthetase molecule, an enzyme which, it should be remembered, has also been shown to act as a translational repressor of the Q β coat protein cistron.

SEMICONDUCTORS

Nothing is the Last Word

from a Correspondent

THOSE were the words of Professor C. A. Coulson (University of Oxford) registering surprise that his description of the vacancy in diamond, developed eleven years ago with Mary Kearsley, should have survived so long. He was making an all-too-brief appearance for the panel discussion on theoretical aspects of defects at the International Conference on Defects in Semiconductors held in Reading from July 19 to 21. Dr A. B. Lidiard (AERE, Harwell) was chairman for the panel discussion, and other participants included Dr G. D. Watkins and Dr R. Messmer (General Electric Company, USA) who are currently enjoying much success in treating defects in semiconductors using Extended Hückel Theory (EHT), a theoretical technique taken over from chemistry. Debate centred on the provision in EHT of only one electron states, that is the strict localization of electrons. Drs Watkins and Messmer were the first to recognize this drawback, and they suggested that where the states fell within the usual bands of allowed energy state delocalization would occur. Professor Coulson thought this in the best chemical tradition.

The conference was the sixth in a series started at Gatlinburg in 1958 when some of the first work at liquid helium temperatures was reported. No satisfactory explanation of that work has yet been advanced, but hopes were raised at Reading that such an explanation might soon be forthcoming. Besides such continuing lines of enquiry, the field has since embraced work on ion implantation and amorphous materials although the organizers were careful only to include contributions in these areas which were strictly linked to a study of radiation damage.

Professor J. W. Corbett (State University of New York) gave the first invited paper on defect production and defect production processes. Having cooperated very closely with Drs Watkins and Messmer, he discussed the possibility of very low energy defect migration and production processes, which appear to follow from the "bond centred" and "split" interstitial configurations, now suggested by EHT, together with the effects of charge state on their stability. It was here that some hope of explaining the work at low temperatures appeared.

Apart from their contributions to the panel discussion Dr A. B. Lidiard and Dr G. D. Watkins gave reviews on defect energies and states, and on vacancies and interstitials in semiconductors, respectively. A review by Dr R. C. Newman (University of Reading) and Dr

R. Peart on the role of defects in diffusion produced the general prescription that most existing work should be looked at again, and closely. Results in ion implantation were discussed by Dr F. L. Vook (Sandia Laboratories, USA) with special reference to the way in which it can elucidate known defect centres, for instance by isotopic substitution. Work on devices was reviewed

by Dr B. L. Gregory (Sandia Laboratories) but unfortunately there were no contributions. Dr R. S. Nelson (AERE, Harwell) presented observations on highly disordered materials using the electron microscope, and the potentialities of helium ion channelling techniques for evaluating disorder were clearly set out by Dr W. L. Brown (Bell Laboratories, USA).

More Evidence for Repressed Messengers

WAY back in 1963, McAuslan noticed that actinomycin D could prolong by many hours the accumulation of virus-specific thymidine kinase in pox-infected HeLa cells. He proposed that during the normal course of infection an RNA was made between 2 and 4 hours, which repressed the expression of the mRNA for TK made before that time.

Soon afterwards, the idea of post-transcriptional repression of preformed mRNA function was taken up again by Gordon Tomkins, also to explain paradoxical effects of actinomycin D on enzyme induction. Hepatoma tissue culture (HTC) cells exposed to adrenal corticosteroids show an increase by a factor of ten to fifteen in the amount and rate of synthesis of tyrosine aminotransferase (TAT), which reaches a steady state with enzyme degradation in about 12 hours. If the steroid is removed at this time there is, within 6 hours, a decline in the amount of enzyme and its rate of synthesis to the basal level. If, however, 5 $\mu\text{g/ml}$ of actinomycin D is given when the steroid is removed, the enzyme level increases ("superinduction") or remains at the same level for many hours. From these and many other observations, Tomkins and his colleagues have devised a model for the control of TAT synthesis which involves two genes—one for the mRNA of TAT itself, and another for a labile RNA or protein repressor molecule which binds to TAT mRNA and both prevents its translation and promotes its degradation (Tomkins *et al.*, *Science*, **166**, 1474; 1969). According to this model steroids could act entirely at the transcriptional level by promoting the synthesis of TAT mRNA and shutting off the synthesis of the labile repressor RNA, or they could work post-transcriptionally by inactivating the repressor and allowing the TAT mRNA to accumulate (provided that it was already being synthesized continuously). High levels of actinomycin D are thought to inhibit the synthesis of the labile repressor, thereby releasing any inactive mRNA-repressor complexes and increasing the rate of synthesis of TAT ("messenger rescue").

Although this model, in one form or another, has been around for a long time, consideration of it has been, in

Tomkins's own words, "rather restrained". In part this has resulted from confusion over the interpretation of the "superinduction" by actinomycin D of TAT activity. Kenney and his collaborators, working with a different hepatoma cell line (H35), under conditions in which TAT was induced in monolayers maintained in Eagle's basal medium without serum, found that enzyme turnover after removal of the steroid was strongly inhibited by high doses of actinomycin D. They therefore argued that "superinduction" was an artefact produced by this inhibition of TAT degradation. Tomkins's reply to this was to show that under these poor nutritional, stepdown, conditions, the degradation of TAT was faster than in an enriched medium, and that while such "enhanced" degradation was indeed inhibited by actinomycin D, turnover under the conditions normally used in his laboratory was not affected by the antibiotic. Instead, he has consistently argued that "superinduction" is a result of an increased rate of synthesis of TAT, as predicted by the post-transcriptional control model, and in an article in *Nature New Biology* next week (September 6) he presents new evidence in favour of this interpretation.

Obviously, until direct evidence for translational repression of TAT messenger function is available, the model must remain hypothetical and open to alternative explanations, however unlikely. (For example, in the presence of high doses of actinomycin D, labile mRNAs will be lost and such things as initiation factors, which may be limiting for protein synthesis, may become available for increased translation of the more stable mRNAs.) Meanwhile, in his article in *Nature New Biology*, Tomkins presents an impressive list of proteins and enzymes like the pox-virus thymidine kinase cited earlier, the concentrations of which have been reported to be stimulated by actinomycin D or other inhibitors of RNA synthesis. Some of these effects may, of course, result from an inhibition of protein degradation, but others, it is argued, may turn out to reveal a widespread use of post-transcriptional regularity elements in the control of the utilization of mRNA in animal cells.

A Year of Transition for Basic Science

JOHN MADDOX

THE year which has passed since the last annual survey of some especially exciting aspects of scientific research has done very little to restore the optimism about the future course of development which lasted for a decade from the mid-1950s. In Britain, after all, the past year has seen the prolonged debate about Lord Rothschild's argument that the work of the research councils should be geared more closely to the needs of government departments, and the essence of his case is now accepted. This followed closely on the recognition by the universities that the pace of growth of expenditure on academic research would not in future keep up with the growth of the university population, faculty and students, with the result that increasing numbers of British academics must look forward to a life of teaching, not research mixed with teaching. In the United States, where the grant-giving machinery that Lord Rothschild asked for has been long established, and universities have grown accustomed to carrying out a good deal of applied research, parallel forces have been felt in basic science. Although the past year has seen a welcome increase in the total amount of the funds available, the National Science Foundation and the Administration have been together flirting with schemes for spending substantial amounts on project-oriented research, while the much-vaunted \$1,000 million programme is likely to be increasingly spent, as the years go by, on applied research. And this is not a tendency confined to Britain and the United States. The government of India, for example, is just as anxious that it should get what is called a return on its money. The European Commission, meanwhile, is eager to find some means of sponsoring productive applied research in Europe, while there is every likelihood that the funds available for universities will grow less quickly than the numbers of people anxious to work in research. What do these trends imply for the years ahead?

The first thing to be said is that it is right and proper that governments should want to see more tangible benefits for what they are prepared to spend on scientific research and development of all kinds. After decades of being told, principally by academic scientists, that science is the mainspring of industrial prosperity, it is not surprising that they should now be asking to see the colour of the money of basic science. Moreover, there is every reason why the scientific community itself should welcome any arrangements for the administration of science that may help to make its craft more useful, and more explicitly so. In retrospect, and as things have turned out, it is, for example, hard to see how it might be possible to justify the vast expenditures on high energy physics which there have been since the Second World War, and whatever the excitement of this important field of science, governments would be well within their rights to complain that they have been misled by the scientific community in the proper allocation of priorities within academic science. It is lucky that politicians have so far been comparatively polite—or even indifferent—on this

issue. But the biggest danger now is that there will be a much more widespread disenchantment with the value of academic science as a whole. That would be a disastrous mistake.

The truth is that the past year, like its predecessors, has been rich in excitement and full of promise, intellectually and in usefulness. The following pages draw attention to only some of the work that has made a mark. Astronomy continues to boom both in experiment and theoretically. The past year has seen the flowering of X-ray astronomy, while arguments like that of Hoyle and Narlikar (*Nature*, **238**, 86; 1972) show that interest is reviving again in Eddington's old problem of how to make an explicit connexion between microscopic physics and the large-scale structure of the universe. Black holes have bobbed on the scene, and will not easily be talked away. In the Earth sciences, the great excitement about continental drift, or more accurately plate tectonics, has now led to the formation of common ground between geophysics and geology, to the common benefit. In much the same way, the recently somewhat abstract molecular biology has served in important ways to fructify other fields of cell biology. The recognition in an actual system of reverse transcriptase, now two years old, has helped enormously to improve the understanding of how viruses infect cells, and the implications are especially exciting for cancer viruses. Sub-cellular techniques have made it possible to separate and characterize specific parts of cell membranes—receptors in nerve cells and hormonal receptors in other cases. Work with the prostaglandins, and especially their relationship with the drug action of aspirin, promises to be of great practical importance (Vane *et al.*, *Nature New Biology*, **238**, 104; 1972) as does that with interferon. In the past year, however, the immunologists have captured most attention, not merely because it is now much easier to characterize the complicated molecules of antigens and antibodies, but because there is a good deal of experimental work to show that the techniques of immunology may be successfully applied to a much wider range of diseases, cancer as well as infections by protozoan organisms, not merely bacteria. And there is just a chance that what has happened in immunology in the past year will turn out to have a bearing on the outstanding problems of biology—those of understanding how the different cells in a single organism acquire different functions. Given that the past year has also seen a start on the problems of molecular genetics in organisms more complicated than bacteria and the beginning of work on the inheritance of behaviour, it is plain that the most urgent question which many academic scientists face is where they should next turn their energies.

So is there a danger that the promise of these exciting developments, in the advancement of science as well as in practical applications, will be undermined by the pressures—real or imagined—to which academic scientists feel themselves exposed? It is first of all important that the

universities have by no means been solely responsible for the developments which have gripped the imagination in the past few years. Indeed, the time is long since past when the universities were dominating sources of innovation in basic science. Increasingly since the Second World War, a great variety of independent research institutes have made important contributions. Some are government or government supported laboratories working more or less independently of universities. Others are private research institutes, within or just outside the walls of some lucky campus. Some are very large, others are tiny by comparison, with mere handfuls of research workers following a single line of inquiry. And although there are important difficulties in the constitution of research laboratories in basic research which are independent of universities, not least of all in the difficulties of keeping people usefully employed during the whole of their careers, there is now ample proof that research institutes are exceedingly efficient places for the conduct of basic research.

The chances are that the financial pressures which now appear unavoidable will accentuate the flight from the universities to the institutes. With the prospect of increasing responsibility for teaching, but especially if there is a serious danger that university academics will be unable to recruit support for the research on which they have set their hearts, there is bound to be a tendency for ambitious young men and women to seek less distracting places at which to carry out their work. Nobody can blame them, but the movement has serious dangers for the health of the scientific enterprise. The common belief that teaching responsibilities stimulate research may have been overdone, but there is no dispute that teaching staffs which are not stimulated by the presence intimately among them of people actively engaged on basic research are likely to become sterile. And can it be an accident that the universities with the highest reputations in research are mostly those at which potential undergraduates compete most vigorously for places? What this implies is that if there is now to be a movement of people with a deep interest in basic research from the universities to research institutes, there must be a compensating attempt to forge closer links between the institutes and the universities. It is in no sense absurd to think that in many cases, research institutes should be deliberately turned into graduate schools for the universities with which they may already be linked. It goes without saying that arrangements like these, which are unlikely to become widespread for several years (common though they may, for practical purposes, already be in the United States) imply significant changes in the character of academic life, even in such mundane matters as the institution of tenure.

If, in the course of years, it were accepted that there must be a change in the way in which research of an academic character is carried out, it would be possible to look for other benefits than mere efficiency. Such a development could, for example, bring about important changes in the channels of communication within science of an academic character. The academic freedom of which universities are rightly jealous is often equated with the freedom to publish, a principle frequently corrupted to the view that an academic is successful only if he publishes a lot. It is true that in present circumstances, the pressure to publish may be just as strong for people working in research institutes—not merely is there a

temptation to keep up with the academics, but institutes like to do what they can to impress their sponsors. Yet everybody knows that informal communication, not merely at scientific meetings but by telephone, is increasingly important as a means of keeping up with quickly moving subjects—in a sense, there has been a partial return to the informality of the 1920s. And research institutes which consist of teams of people with some degree of permanence, and in which the senior people are not outnumbered by graduate students, are at once more able to keep closely in touch with similar laboratories elsewhere and also to evaluate the importance of what they and their colleagues have accomplished. Moreover, in such a setting, there is at least a hope that claims for priority will not always depend on evidence of the first formal publication—an uncertain test at the best of times. So is it unreasonable to hope that the strengthening of research institutes that may be encouraged by the present pressures on basic science may reduce the pressure to publish? It would be a splendid thing if the need for urgent publication were confined to circumstances in which news of some innovation must be communicated to people who do not know already, people working in closely related fields, for example.

The potential benefits of such arrangements are not the only likely consequences of change. Perhaps the most important defect of the organization of academic science in research institutes is that those who work in them are frequently less free to speak out than academics pure and simple. This is especially true of those issues on which the community at large must rely on scientists for informed comment on matters not strictly of a specialized character. Institute workers are often employed directly by governments or by para-government institutions, for example, which implies that it is frequently impossible for them to make public views on the way in which government science is organized. Often there are the susceptibilities of private sponsors to be considered, while many institute directors hold to the mistaken belief that their institute's reputation will be compromised if its members are seen to have views about such things as supersonic aircraft or disarmament negotiations. Moreover, institutes vary enormously in the freedom with which they will allow members of their staff to travel to scientific conferences.

At a time when the public view is, rightly, that professional science should be much more constructive in the formation of public policy, it would be most unfortunate if the likely tendency of the next few years were to diminish the freedom to play an active part in the formation of a public view on scientific matters. If anything, there is a case for asking that the freedom of those who work in existing institutes for research to engage in this important task should be quickly liberalized. But how to create the essence of academic freedom on extra-mural matters within the framework of research institutes? The most urgent need is to ensure that directors are, as they should be, solely responsible for the conditions under which they work, and to ensure that research workers know that it is a part of their duty to lift their eyes above the bench much more often than they are now inclined to do. It is true that such an arrangement might frequently lead to circumstances in which a government employee was found to be helping to inform public opinion on the wisdom of some government decision, but this is not, at the present time, an offence against reason.

A Corporate Conscience for the Scientific Community ?

PAUL SIEGHART

173/176 Sloane Street, London, SW1

An interdisciplinary working party¹ has been trying to see whether the diffuse debate about the social responsibility of scientists can be brought into sharper focus and whether some practical means could be devised—within the framework of the existing social system—for scientists to discharge the special social obligations which many of them feel they have.

Our paper took us two years to write, and is too long to be printed here in full. This article sets out the gist, but without much of the supporting argument.

The first thing to say is that we came to no radical conclusions. Those which have been proposed fall into three broad groups. First, that there should be something in the nature of a Hippocratic Oath or code of ethics for all scientists², whereby they bind themselves not to take part in work which will have socially harmful consequences; second, given that we have an increasingly science-based civilization, that moral and political decisions about the social application of scientific work should be left increasingly to scientists themselves; and third, that there should be radical—if not revolutionary—reform of the whole social system. Of these, we think the first impracticable, the second dangerous, and the third beyond our competence.

(1) The field of science is vast, and the effects of the work done by the people who operate in it are complex and often quite unpredictable. There is nothing inherently good or bad about any advance in the scientific field: the consequences for society will depend on the use which society makes of it. Because of the diversity and complexity of possible social effects, any attempt to draw up a set of general ethical rules for scientists is bound in the end to reduce to a series of pious generalities, incapable of giving a particular scientist any help or guidance in a particular situation. However much the idea may appeal to the popular imagination, we do not consider that the drafting of such a set of platitudes would give any practical help in solving our problem.

(2) We suspect that the idea that scientists should be left to decide how society should use their work is based on a tacit confusion. If decisions of this kind were "scientific" decisions, there might be something to be said for leaving them to scientists. But they are not: they are essentially political and

moral decisions. We have no evidence to suggest that scientists are better qualified to take decisions of this kind than men trained in other disciplines, or indeed that scientists as a class are more rational, or more altruistic, or in any sense morally "better" than other classes of human beings. Neither their training, nor their experience in laboratories or academic institutions, particularly equip them to make important decisions for the community at large. Nor has past experience with different forms of government shown that our society would be better ruled by a small élite of specialists, not subject to the checks and balances designed to avoid too great a concentration of power among too few people, and not accountable to the public whom they serve.

(3) Many people today think that there is a great deal wrong with our social system. Some go on to say that nothing short of a revolutionary change in its fundamental structure will suffice to put things right. In the light of past history, we ourselves do not believe that revolution in Great Britain is likely in the long run to be more effective than reform in improving the quality of life enjoyed by most people. Besides, we have deep-seated objections to violence. For our part, therefore, we start from the premise that "The System" is with us, and will be with us for some time to come, but that there may be some ways of making it work a little better, and improving the prospects of the species which has adopted it.

Our concern is therefore to argue the merits of some such ways, and to suggest one in particular which we think might help our society to make better thought-out and better informed choices, and at the same time give scientists a greater opportunity of putting their social consciences to work for the common good.

Shortening Time Scale

The response time available to society for successful adjustment to new techniques seems to be shortening rapidly. There was less time for adapting to an industrial society than there was for adapting to a pastoral or agrarian one. We have only had about two generations to adapt to the motor car, and less than one to adapt to the jet aeroplane and the nuclear weapon. As the available response time shortens further, it becomes increasingly difficult to think out the best response thoroughly enough, to contemplate all the consequences of different ways of responding, to try out different ideas in order to see whether they work, to make plans in the leisurely atmosphere in which good plans should be made—and, above all, to think, talk, and think again for long enough to ensure that everyone likely to be affected (and increasingly often that means all of us) has had his say, and that his say has had a fair chance of swaying the communal decision.

More and more often, decisions have to be taken too late, when there is little time to think them out fully, to discuss them publicly, to have second and even third thoughts, and to hear what people with different points of view have to say. In the result, often nothing is done at all where clearly something should have been done, or else what is done is a hasty measure which has to be altered later, by which time it may be too late to do what should have been done in the first place.

In our view, the greatest scope for influencing the future of our species favourably lies in a more efficient use of the response time available to society for deciding how new work in the

¹ The members of the working party were: Dr B. S. Drasar, bacteriologist, St Mary's Hospital Medical School, London; J. C. B. Glover, philosopher, New College, Oxford; Dr V. A. S. Glover, neurochemist, University College, London; Dr M. J. Hill, biochemist, St Mary's Hospital Medical School, London; Dr J. Issroff, psychiatrist, Tavistock Institute for Human Relations; D. A. Parfit, philosopher, All Souls College, Oxford; Paul Sieghart, retired barrister (convenor).

² In our use of words like "science", "scientist", or "scientific research", we mean to cover the whole field of science and technology in all its applications, the whole of research whether pure or applied, and all the people—whatever they may call themselves—who work in it.

field of science and technology should be used. That is therefore the area on which we have concentrated our attention.

It is in this area also that the greatest opportunities arise for scientists themselves to discharge the moral responsibility to society about which they are becoming increasingly concerned. But what does that responsibility amount to? To what extent, if any, has a scientist, as a scientist, special moral obligations to society, over and above those which he owes in his ordinary capacity as a citizen?

Special Obligations

A good starting point for deciding whether anyone has a special moral obligation to others is to ask whether he is especially well placed to benefit or harm them. This is, after all, why we say that motorists should drive with care, that parents should take an interest in their children's progress at school, and that doctors should obey their professional code. There are many other groups which one could mention: in each case, the people concerned are in a special position to affect the welfare of others.

The outcome of scientific work can often have a great impact, for good or ill, on other people. Quite frequently, scientists can predict this outcome earlier, and more accurately, than others. Sometimes they can even modify the results. One could claim, therefore, that scientists are in one of those special positions which give rise to special obligations.

In our paper we then consider, in detail, some possible objections to this claim, such as "much scientific work has no effect on society"; or "the social effects of a piece of scientific work may not be predictable"; or "the scientist only provides the knowledge and the techniques: it is society's responsibility how they are used"; or "if I don't do it, someone else will"; or "as a scientist, I am only concerned to increase knowledge. That must take priority over everything else".

After a full analysis of their merits, we conclude that none of these objections is well founded, and that scientists do have special obligations to society. The next question is what these obligations are. The principal ones which have been advanced by others can be conveniently, though loosely, classified as follows: first, to refrain from doing some kinds of work; second, to do some kinds of work; third, to contribute to a greater sense of social responsibility among other scientists; fourth, to think out the social consequences of their own work and of other work in their field; and fifth, to inform the public about that is being done, and its likely consequences.

We consider each of these separately, and conclude that they do form part of a scientist's special obligations. In particular, we believe that anyone who engages in an activity which might have harmful social consequences ought to apply his mind to those consequences, and to act on his conclusions. We do not think that it is good enough to put on blinkers, concentrate on the job in hand, and leave the consequences to others. Although it is easy to state this obligation in general terms, its application in particular cases will often be difficult. The social consequences of a particular piece of research or application may be far from obvious, and their evaluation will often lie outside the competence of the scientist concerned. But if it appears that work which is being done in a particular field might have a significant social impact, it is in our view of the utmost importance that society at large should be given warning in advance. Sometimes, it may be obvious at an early stage that a particular discovery is likely to have clearly harmful social consequences. In that situation, the rest of society should have the opportunity to begin as soon as possible to work out a response which will minimize or eliminate them. On other occasions, important beneficial consequences may be foreseen. Here society should have the opportunity, as soon as possible, to reinforce the research effort, and to prepare the means for utilizing the benefits and making them available to its members. More often, it may be clear that there will be some consequences, but not clear whether they will be harmful, or

beneficial, or both. Here society must have time to work out an intelligent response in advance, by initiating whatever research is necessary to get a clearer idea of what the consequences are likely to be, and what needs to be done about them.

No One's Job

We next consider how the scientist's special obligations can best be performed. Of the five classes of social obligation which we have listed, the first two (to refuse to do, or to do, certain kinds of work) are matters for the consciences of individual scientists, which they can only decide in accordance with their personal moral values. The other three classes of obligation differ from the first two because they can be performed in an organized fashion, and the most important conclusion to which we have come is that, in Britain at least, the institutional machinery available to scientists for the most effective performance of their special social obligations is at present quite inadequate. No doubt this is one of the principal reasons for the current sense of ethical frustration in the British scientific community, and the mounting spate of discussion on ethical questions within it.

Having rejected the idea that the scientific community should itself take decisions about the social consequences of its work—which would put it into a position above national governments, where it was answerable to no one but itself—it follows that such decisions, like all others affecting the community at large, must remain "political" decisions, and must be taken by the organs of government which the community has chosen for itself (or, in some countries less fortunate than ours, has had imposed on it). In Britain, that means parliament and the government of the day, advised by the permanent officials of the departments.

What particularly concerns us is that, at present, there appears to be far too little communication between the organs of government on the one hand, and the scientific community on the other. We do not doubt that the government of the day has at its disposal the best scientific advice if it chooses to call for it, but it is not the primary function of those advisers to press on politicians and civil servants unasked-for views, nor to stimulate discussion and planning about possible dangers, at all events before they have become a matter of public concern. Correspondingly, politicians and their permanent advisers are busy men, and it takes most of their time to deal with those matters which are of immediate urgency. It is not their function to find out what is going on in remote laboratories, and to think about the possible beneficial or harmful consequences to society of the research work which is being done there. The trouble is that, at the moment, there is no single institution in this country whose chief job this is. We think it is high time that it should be someone's job.

The present position is too often chaotic. Scientific work which may have a great impact on all our lives is being done all over the world. The scientists who are doing the work may well be aware, even if sometimes only dimly, of the possible dangers, but it is no part of their task to work out the appropriate social response, nor are they qualified to do this. To say, as we do, that they have an obligation to inform the public about that they are doing, and its implications, is easy enough. But at the moment there is little the individual scientist can do in this respect between the extremes of a sensational newspaper article, and a discreet letter to a member of parliament. It is not until the potential dangers become apparent through some scandal or disaster that they catch the attention of the popular press, and there is a public outcry. At that point, politicians and their advisers begin to take an active interest and look—often frantically—for the appropriate response. But by that time, only too often, it is in one way or another too late, for too many people or organizations have acquired a vested interest in the discovery, or its development, to make it possible to deprive them of its fruits. The response will then be too little and too late, and we either have to put up with the harmful

consequences with little or no mitigation, or we are subjected to the constraints of panic measures which are ill-thought-out and which later prove to be inadequate or to have unexpected side effects.

A New Institution

We think that this situation could be much improved if there were to be brought into existence a body, organized by the scientific community itself and expressly charged with the task of informing the public in general, and the organs of government in particular, at the earliest possible time, of all scientific work likely to have important social consequences for good or ill. This would serve to lengthen the available response time, and eliminate the present lack of communication. Also, individual scientists would have an opportunity to perform their social obligations in many situations where, as things now are, they can do nothing.

We therefore propose the setting up of a new institution, which for the sake of brevity we simply call "the council". This should be an organized body of scientists, whose task it is to stimulate informed public discussion about the possible consequences of socially important pieces of scientific research, in each case at the earliest possible moment. The most important thing about such a body is that its voice should not go unheard. It must therefore be a responsible body, composed of scientists (and perhaps also some others) whose views will command respect not only among their colleagues, but among politicians, civil servants, editors and the general public. We are fortunate in having a number of men and women of this calibre in Britain, and we feel sure that enough of them would be willing to give time to the performance of such an important social duty.

But no one can expect scientists of the necessary level of distinction to do all the administrative work which will be required. The council must therefore have a small, but highly competent, full-time secretariat. Beginning perhaps with one able, hard working and well qualified secretary, with a staff of one or two, this could be enlarged later as circumstances require. We envisage that in practice the secretariat would take most of the initiatives, and would stand in much the same relation to the council as the permanent officials of a government department stand to their minister.

We regard it as central to the function of the council that it should be, and be seen to be, wholly independent of any kind of political or financial influence, from any quarter. We do not therefore think that it ought to be set up or sponsored by the government of the day, or even by parliament. We do not mean to suggest that, if it were, either of these organs would in fact bring pressure to bear on it. But if it is to enjoy the public credibility and trust essential to the proper performance of its task, what is important is that the public should have no cause to suspect that pressure of this kind could be applied. Besides, there are several bodies—such as the various research councils—which are sponsored by the government and which perform important functions of their own. It is not our intention that the council should in any way overlap with, or duplicate, any of these. If, therefore, it needs any sponsoring authority, that should be the scientific community itself, acting through one or more of its existing organizations.

What It Would Do

We envisage that the council would have ready access to expert scientific opinion in all fields falling within its domain. If it wishes, in any given case, to entrust a piece of thinking, or the drafting of a report, to a working party of experts, we feel sure that it will find ample talent to draw on within the scientific community. It will almost certainly need to add to such working parties—as we have done ourselves—members from disciplines other than those of science, and here again we feel sure that there are enough people with the necessary qualifications, and with sufficient concern for the social effects of science, to be willing to serve.

There remains the question of finance. The members of the council themselves, and any experts they wish to consult or who serve on its working parties, will presumably give their services gratuitously, in accordance with the long-standing tradition which obtains in Britain in matters of this kind. But this cannot be expected from the secretariat, and from time to time the council may itself wish to carry out research which has to be paid for. The necessary budget will be minute in comparison with the social value of the work done, but the money still needs to be found. For the reasons we have already mentioned, we do not think that it should come from the government. The obvious source would be those institutions which themselves are representative of the scientific community, and perhaps one or more of the charitable foundations who, in so many fields of this kind, make possible work of the utmost value to the community which could not be done at all without their generosity.

The creation of the council will, we hope, not go unnoticed by the scientific community. It will probably advertise its existence in the scientific press, and the secretary may write to university departments, research laboratories and even individual scientists, to invite their contribution to its work.

That contribution, as we see it, should in the first instance consist in the reporting to the council of any work on which a scientist is engaged and which in his view—or that of others whose opinions he thinks worthy of respect—is likely to have social consequences for good or ill. It is in this way that the individual scientists can most easily, and most effectively, perform many of the special obligations which we think he owes to society at large. Such a course of action has many advantages for him, since it relieves him from having to work out for himself, often with insufficient resources, the possible consequences of his work, let alone what would be the appropriate social response. In effect, once he has reported his concern, the rest is up to the council in its capacity as the working conscience of the scientific community as a whole.

In response to its invitations, the secretariat will probably receive a good number of reports. Some may need further clarification with the scientist concerned. With others it may be obvious that they are of no immediate urgency. The odd one may be so clearly fanciful that it can be safely filed. But it is likely that quite a number will need urgent and detailed attention. We see the secretariat bringing these forward to the council itself—or perhaps to an executive committee—for decisions on their order of priority, and on what needs to be done about them.

Practical Problems

In our paper we then give an illustration of how the council could work in practice, choosing as our example research on the predetermination of the sex of children. An interdisciplinary working party appointed by the council studies the problem, commissions a social survey, and drafts a report. This is published in the council's name; as the matter is one of great public interest, the press and the other media can be relied on to give it ample publicity. From that point on, the normal democratic process can take over. The matter will become one of continuing discussion, in the press, in parliament, and in the departments of government most concerned. By the processes with which we are all familiar, the response thought to be most apt for our particular society can evolve, and eventually any necessary legislation can be passed.

The important point for our purpose is that, as a result of the council's work, the public debate and the political decision-making process will be based on a better knowledge of the relevant facts, begin much earlier, and be taken at a more leisurely pace, and will therefore have more chance to lead to a policy that is properly thought out. Perhaps even more important, all this will happen before any vested interests have been created which it would be unjust or impolitic to disturb.

The first couples will not yet have chosen the sex of their children and so find themselves, and their private affairs, on the front pages of the newspapers; no drug company will yet have spent large sums on research, development and the creation of production capacity, and the scientist concerned will receive praise for his early warning rather than blame for having made yet another contribution to the deterioration of the quality of human life.

This hypothetical sketch serves to show how the council might choose to perform this part of its function. But it could later decide to extend its functions in several ways. It might for instance choose to start enquiries into fields of research likely to have a social impact, rather than to wait until an individual scientist sends in some information. Or it might recommend, or itself initiate, a piece of social research. Sooner or later, it is bound to receive suggestions from scientists and others for socially desirable scientific projects: if it thinks them sufficiently important, it may well decide to say so, without, of course, trespassing on the functions of other bodies which already exist and which perform adjacent tasks.

It is, of course, not our intention that the council should make political or moral pronouncements, or public value judgments on contentious issues. That is not its function. Its reports will rather be in the contingent form: "If the present work in this field is successful, then the following social consequences seem probable or possible. . . . Should these consequences be thought undesirable, then the following counter-measures could be considered, and if they were used they would have the following probable effects. . . ."

It remains to deal with one problem, which we recognize as being one of substance, and that is the problem of confidentiality. Although no reputable scientist—outside the military field—would undertake research on the terms that he would

not be allowed to publish his results, it is not uncommon for him to accept a restraint on publication until his industrial employer had had the opportunity of applying for a patent, or otherwise obtaining a lead over his competitors. In such a case, the sending of a report to the council might be a breach of that restraint. This could sometimes create a delicate situation, but we do not think that this problem is grave enough to outweigh the advantages which our proposal would hold out to the community. In many cases, it should be possible for the council to arrive at a negotiated solution with the commercial interests concerned. In any case we would hope that, once the council was well established and commanded the public respect which we anticipate, it would become an accepted practice among scientists to negotiate the necessary exemption from confidentiality in this one respect as part of their terms of employment. No such problem faces those scientists who choose to serve the military: they are bound by the Official Secrets Acts, and would expose themselves to criminal prosecution were they to communicate their worries to the council without the consent of their superiors. To them, therefore, the use of the proposed institutional machinery would be denied and they would be left, as they are now, to settle such problems with their private consciences.

As a final objection, it may be said that a body designed to articulate the social conscience of British scientists is of little use in a world where science, and its effects, are international. We answer, first, that the importance of British science in the world is, even today, disproportionately high in relation to Britain's population; and, second, that if this is a good idea, the fact that other countries have not yet adopted it is not a valid objection to putting it into practice ourselves. If it is done here, and is seen to work, it will not be long before it is taken up elsewhere too.

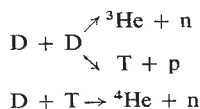
The Road to Controlled Nuclear Fusion

L. ARTSIMOVICH

Department of General Physics and Astronomy of the USSR Academy of Sciences, Leninskiy Prospekt 14, Moscow B-71, USSR

Professor Artsimovich describes the present state of research into controlled thermonuclear fusion and assesses its prospects.

THE problem of controlled fusion lies in the technical exploitation of the energy contained in the lightest elements, particularly the isotopes of hydrogen, deuterium (D) and tritium (T). This energy is released in the fusion reactions



where n and p signify neutrons and protons respectively.

The energy released in these reactions per unit mass of nuclear fuel is several times larger than that released in the fission reaction of uranium, on which contemporary nuclear power depends. If allowance is made for the possibility of making use of the T and ${}^3\text{He}$ formed, the energy available from the combustion of 1 g of deuterium is about 100,000 kWh.

The exploitation of fusion reactions as a source of energy is a highly desirable goal for two chief reasons. First, the primary nuclear fuel—deuterium—constitutes a practically limitless source of energy, because it can be extracted directly from water, in which there is one atom of deuterium for

every 7,000 atoms of ordinary hydrogen. Second, nuclear fusion does not generate as much radioactive waste as nuclear fission, so storage—a serious drawback of conventional nuclear plants—does not constitute such a serious problem. By contrast with nuclear fission, however, nuclear fusion reactions cannot be sustained by themselves in matter which is in its normal state. To induce intense nuclear fusion reactions deuterium, or a mixture of deuterium with tritium, has to be heated to a temperature of several tens of millions of degrees. Only under these circumstances do the atomic nuclei have a velocity sufficient to permit collisions in which the particles have a high probability of overcoming their mutual electrical repulsion and initiating a thermonuclear reaction. In the temperature range at which this begins to occur, matter can exist only in a state of full ionization, that is, as a hot plasma consisting of a mixture of rapidly and randomly moving atomic nuclei and electrons. A high temperature deuterium (or deuterium-tritium) plasma is the new nuclear fuel which will fire the boilers of the future thermonuclear generators producing electricity from nuclear fusion.

Temperature and Confinement

In order that this new nuclear fuel can play its role, however, two conditions have to be fulfilled. First, the temperature of the ions (nuclei) of the plasma must exceed a certain minimum value, such that the energy released in the nuclear fusion reactions exceeds the energy lost from the plasma by X-radiation (resulting from collisions between ions and

electrons). This temperature threshold lies, for a pure deuterium plasma, at around 200 million K, whereas for a pure deuterium-tritium mixture with equal concentrations of both components this is reduced to about 50 million K. But even comparatively insignificant concentrations of heavy atoms lead to a rapid increase in the threshold value of the ion temperature.

Second, the thermal energy has to be confined in the plasma for a sufficiently long time. The necessary condition which has to be fulfilled is $n\tau > A$ where n is the density of the plasma (the number of electrons cm^{-3}), τ is the average confinement time of the energy (the ratio of the energy content of the plasma to the thermal flux from it) and A is a constant depending on the composition of the nuclear fuel.

The meaning of these conditions is easily explained. After "ignition" the thermonuclear fuel will proceed to release its contained energy only if each atomic nucleus in the plasma has a sufficiently high probability of undergoing a nuclear collision leading to a fusion reaction in the time τ . This probability is proportional to τ and to the density of the plasma and consequently it is necessary that their product should exceed a certain limit. For pure deuterium the minimum value of $n\tau$ is $\sim 10^{16}$ and for a mixture of equal parts of deuterium and tritium it is significantly lower—about 10^{14} . I shall not attempt to give more precise numerical values of the constant A as it depends on the design of the thermonuclear reactor.

It follows that a mixture of deuterium and tritium is a much more effective thermonuclear fuel than pure deuterium. Tritium is almost non-existent in nature, however, and only comparatively small quantities of it can be prepared with the help of ordinary nuclear reactors working with uranium.

Consequently tritium can only be used in thermonuclear power generation if a method of producing it in bulk within the thermonuclear reactor itself can be developed. In principle this process is possible, if use is made of neutron multiplication reactions (reactions of the $n \rightarrow 2n$ type). I shall not discuss this process here, except to observe that deuterium alone will probably eventually play the role of the new nuclear fuel (if one takes a sufficiently long term view of thermonuclear power).

It is already more than twenty years since the USSR, the United States and Britain made the first attempts to produce in the laboratory this new nuclear fuel—a deuterium plasma. Since then many scientific organizations in other countries have taken up the problem of controlled thermonuclear fusion. These investigations have widened considerably and have acquired the character of fundamental scientific research. But the problem, which initially had a very optimistic prognosis, now seems so difficult that research is still at the level of basic physics, and the achievement of the practical goal is still very far off. The principal difficulty at present is still the creation in the laboratory of methods of producing a high temperature plasma with parameters which correspond to the requirements described above. It has not yet proved possible to fulfil simultaneously both of these requirements in any one experimental apparatus. Nevertheless the ultimate goal is being approached, by degrees, and in the process several different ideas have been pursued.

Magnetic Confinement

Most methods used so far to create a high temperature plasma are based on the idea of magnetic confinement (or magnetic thermal insulation); the plasma is trapped in a very strong magnetic field. The geometric form of the magnetic field is chosen such that the electrons and ions of the plasma, which can move comparatively freely only along the lines of force, never escape from a finite region of space within which the plasma is to be contained. Under these circumstances the transport of energy by fast electrons and ions to the walls of the chamber within which the plasma

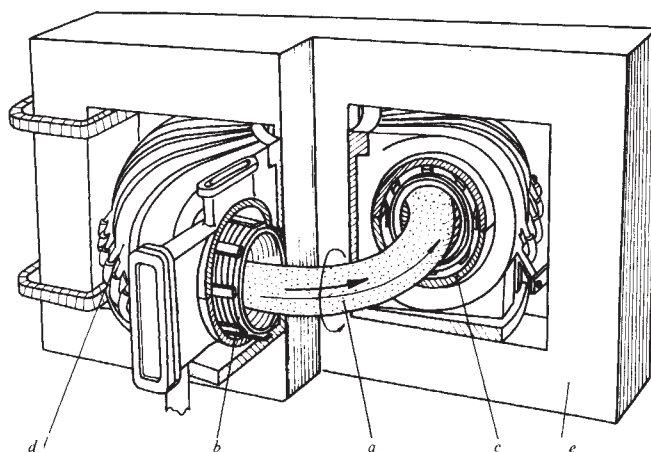


Fig. 1 Schematic representation of a Tokamak device (sectionalized). *a*, The plasma column; *b*, the inner vacuum chamber (thin stainless steel); *c*, thick copper shell; *d*, the windings which generate the external longitudinal magnetic field; *e*, iron core of the transformer. Arrows indicate the current in the plasma and components of the magnetic field.

is situated can be reduced to an extremely low value, that is, the heat is well confined within the plasma. The principle of magnetic confinement of high temperature plasma was suggested more than twenty years ago and, incidentally, was discovered completely independently by physicists in the USSR, the United States and Britain. It has served as a point of departure for research on controlled thermonuclear fusion and as a source for a number of basic lines of investigation in this field, which can by now be described as traditional.

In recent years the number of lines of investigation has increased as a consequence of the arrival of some new suggestions—for the extremely rapid heating of matter by laser pulses or by electron beams of very high power. In such circumstances the use of magnetic heat insulation can become unnecessary. It is not yet possible to make a final choice between the various approaches to the solution of the problem of controlled fusion, and it is still early even to indicate the most promising direction along which one ought to concentrate the chief line of attack. In this field there is no royal road. Consequently—let "a hundred flowers blossom"—at least for several years to come the programme of scientific investigation must remain sufficiently widely based. In looking at the present situation in the various branches of this general programme, I shall consider first the results obtained in specific implementations of the various concepts of magnetic plasma confinement.

Among these, an important place belongs to experiments for obtaining high temperature plasma in so-called closed magnetic systems. In such a system the plasma can flow freely along the lines of force of the confining magnetic field, which never leave a finite region of space. The plasma is formed within a toroidal vessel and takes the form of a ring shaped loop. In this variant of magnetic confinement the particles of the plasma are, so to speak, disturbed as little as possible. Most of the electrons and ions of the plasma are limited only in their motion perpendicular to the magnetic field, whereas along the field they can move completely freely. These are the so-called "passing" particles. In any closed system, however, there is also a group of "trapped" particles which are forced to oscillate within a finite section of the magnetic field. As they move along the field lines, the trapped particles are reflected by regions of relatively strong magnetic field strength. The trapped particles play an important role in the energy balance of a high temperature plasma, for their

presence increases the thermal conductivity and diffusion velocity within the plasma column.

Two very typical examples of closed systems are devices belonging to the Tokamak and Stellarator families. Tokamak is the simplest system from a geometric point of view for the production of a hot plasma under quasi-stationary conditions. In devices of this type, a current created by induction flows along the ring shaped plasma column, and to this end the toroidal vessel surrounds the iron core of a transformer. The magnetic field of the current plays a leading role in the thermal insulation and confinement in the plasma. The heating of the plasma is due to the Joule heat released by the current. To prevent the development of destructive instabilities, which are characteristic of a plasma conductor carrying a strong current, a longitudinal magnetic field is used. This is produced by external windings, the magnetic field lines of which are parallel to the current. The strength of this field H_0 must be many times greater than the strength of the field of the current H_ϕ . The longitudinal field is generated by means of windings distributed on the surface of the toroidal vessel as shown in Fig. 1, which is a schematic representation of a Tokamak device. The magnetic field within the plasma column consists of the super-position of the two fields and consequently its magnetic field lines have a helical structure. The helix is extensively drawn out along the column, as H_0 is much greater than H_ϕ . The characteristic feature of Tokamaks is the axial symmetry of the magnetic field, which simplifies the production of a plasma column with a large perpendicular cross sectional area.

Studying Tokamaks

The investigation of the properties of plasma in Tokamak devices was begun in the Institute of Atomic Energy in Moscow and has been in progress for more than fifteen years. In the summer of 1969 a group of English physicists from the Culham Laboratory played an active part in these investigations, applying for the first time the laser method of measuring the electron temperature of the plasma to the largest of the Soviet Tokamak devices. Beginning in 1970 the circle of investigators of the properties of hot plasma in systems of the Tokamak type has also included laboratories in the United States, France and Germany. In the study of the physical processes that occur in Tokamak devices, theory and experiment have always gone hand in hand. The choice of specific experimental programmes has been simplified by the fact that even at the very earliest stage in the investigations V. D. Shafranov developed a theory of the equilibrium of a plasma column and introduced the basic rule which ensures its stability with respect to the most dangerous large scale deformations (the so-called Kruskal-Shafranov criterion). According to this rule, the column is

stable if the condition $\frac{H_0}{H_\phi(a)} \cdot \frac{a}{R} > q$ is fulfilled where H_0

is the strength of the field due to the current at the surface of the plasma, a is the radius of the perpendicular cross section of the column—assuming that it has a nearly circular form—and R is the major radius of the toroidal system (the radius of the plasma ring). The quantity q is called the margin of stability: it depends on the distribution of the current across the cross section of the plasma, and if this distribution is bell shaped q has to exceed 2 or 3.

A strong incentive for further experimental work has been provided recently by the development of the "neoclassical" theory of heat transfer and diffusion in toroidal plasma systems. The basis of this theory was laid down by Sagdeev and Galeev; they first showed that the presence of trapped particles whose motion involves significant displacements from the given line of force leads to a sharp increase in the rate of heat transfer in the plasma. The appearance of this new theory substantially reduced the discrepancy which had

previously existed between experimental data and theoretical calculations.

In the Tokamak family the largest device at present is T4. The diameter of the toroidal chamber (2R) in T4 is 1.8 m and the diameter of the cross section of the plasma column can be as large as 35 cm. The external walls of the chamber are made of thick copper sheet. In those investigations which have been carried out so far on T4, the maximum strength of the longitudinal field created by the windings as measured at the toroidal axis of the plasma column has been 40 kgauss and the current in the plasma has reached 230 kA. The lifetime of the plasma column is determined by the time for which the current is maintained in the plasma. In T4 this is usually 0.1 s.

Typically, experiments on Tokamaks are carried out with an average value of the electron density $\bar{n}_e$ in the plasma column of 2 to 3×10^{12} up to 3 to 4×10^{13} cm $^{-3}$.

Results of Experiments

Measurement of the thermal energy of the plasma carried out by a number of independent methods has shown that when the thermal regime has been set up, the ratio of the average value of the kinetic pressure of the plasma $\bar{p}$ to the pressure of the magnetic field of the current at the surface of the plasma column $H_\phi^2/8\pi$ is comparatively insensitive to relatively wide variations in the initial conditions of the process. Except at low densities, this ratio is $\sim 0.45 \pm 0.1$; this result can also be written in the following form

$$N(\bar{T}_e + \bar{T}_i) \simeq 1.6 \times 10^{13} J_A^2$$

where N is the total number of electrons per cm of length of the plasma column ($N = \pi a^2 \bar{n}_e$), $\bar{T}_e$ and $\bar{T}_i$ signify the average values of the electron and ion temperature (calculated per electron or ion in the plasma), and J_A is the current measured in A. The electron temperature usually significantly exceeds the ion temperature, and the maximum value of T_e on the axis of the plasma column in T4 reaches $\sim 3 \times 10^7$ K when $n_e \simeq 3.5 \times 10^{13}$.

Experimental data obtained from measurements of the ion temperature are in satisfactory agreement with the predictions of the neoclassical theory of thermal processes in a plasma column. The maximum value of T_i on the toroidal axis of the column of T4 is $\sim 7.5 \times 10^6$ K for a deuterium plasma. In this case the plasma becomes a source of neutron radiation which has a quasi-stationary character for several tens of ms. It appears that this radiation is of thermonuclear origin, at least when the plasma density is relatively high.

The mechanism of heat loss from the plasma column has not yet been finally established. From measurements of the energy balance it follows that the electron component of the plasma loses its energy very much faster than is predicted by neoclassical theory. The optimum energy confinement time τ_E in the plasma for the experiment T4 is no greater than 10 to 15 ms, but the product $n\tau_E$ determining the effective thermal insulation of the plasma now reaches $\sim 0.5 \times 10^{12}$. The values of the basic physical plasma parameters (T_i and $n_e\tau_E$) which have been achieved so far in Tokamaks are still much smaller than those at which the future thermonuclear reactor might operate. But in the past decade the ion temperature and the energy confinement time in Tokamak devices have undergone a steady increase. In the seven years since the Third Geneva Conference on the Peaceful Uses of Atomic Energy, for example, T_i and $n\tau_E$ have grown by an order of magnitude. One might hope that in the next five years it will be possible, with the more powerful Tokamak devices, to raise both of these parameters by at least a factor of five. This will be achieved if present views on the behaviour of plasma in Tokamaks remain applicable to systems with a higher temperature, in which collisions between particles become very rare. The prognosis may become even more optimistic if there is significant progress in the development

of new methods of heating plasma in Tokamaks for, as the temperature rises, the Joule heating due to the passage of the current through the plasma becomes less and less effective as a source of energy. And as a means of improving the physical parameters of the plasma, it may in future prove useful to change the form of the perpendicular cross section of the plasma column by drawing it out along the principal axis of the system. It is not out of the question that under these conditions one would be able to increase significantly the magnitude of the field of the current H_z , and consequently also the magnitude of the plasma pressure, without varying H_θ . In principle one can think of this as a "secret reserve" of Tokamaks, the tapping of which would constitute a big step forward.

Stellarator

In the Stellarator, another type of closed toroidal system, the magnetic field confining the plasma column is a longitudinal field of external origin, which has a rather special and comparatively complicated structure. The external magnetic field in the Stellarator is characterized by a so-called "rotational transform", in other words the field lines of the magnetic field rotate about the toroidal axis of the plasma column as one moves along the chamber. (The magnetic field in Tokamak is similar, but in that case the helical structure of the field lines arises from the superposition of H_θ and H_z).

In Stellarators the presence of a ring current in the plasma is not essential; consequently the plasma in such systems can be created not only by inducing a discharge but also by ionizing gas with high frequency electric fields or by direct injection of plasma from outside.

The Stellarator programme began at Princeton Laboratory in the United States. Since then experiments of this type have been developed in Britain, Germany and the USSR. Work on Stellarators has had some setbacks, but the indications are that in the comparatively near future Stellarators will also be used successfully to produce hot plasma with large values of the parameters T_i and $n\tau_E$. It is possible that in the process it will be necessary to change somewhat the original geometrical proportions of devices of this type in order to increase as much as possible the perpendicular cross sectional area of the plasma column and to increase the magnitude of the perpendicular component of the magnetic field which is responsible for the thermal insulation of the plasma.

Magnetic Mirrors

Another traditional branch of the thermonuclear programme has been the development of methods of producing hot plasma in so-called adiabatic magnetic traps. In such systems the rapidly moving charged particles of the plasma are confined as a result of reflexion by "magnetic mirrors"—regions with increased strength field in which the lines of force are bunched together. The simplest means of creating such a magnetic configuration is to line up on a single axis two identical coils carrying currents which flow parallel to each other. The strength of the magnetic field within each coil is stronger than in the space between them and the magnetic field lines within the coils are thus bunched together. The regions within the coils function as magnetic mirrors for charged particles: electrons or ions trapped in such a field oscillate between the mirrors along the lines of magnetic force. In adiabatic traps the plasma contains only those particles which are trapped in this way. The "trapping" condition is that the velocity vector of the particles must have a sufficiently large angle with respect to the direction of the field lines. If the velocity vector of a given particle is deflected as a result of its collision with another particle and in consequence makes a rather small angle with the

direction of the magnetic field, then the particle can pass freely through the magnetic mirrors and is lost. This is the principal difference between the behaviour of trapped particles in adiabatic traps and in closed systems. If in a Tokamak, for example, the direction of the velocity vector of a trapped particle is deflected as a result of a collision with some other particle in the plasma, so that it moves at a small angle with respect to the field line, then this simply signifies its transition to the class of "passing" particles, which move freely along field lines and remain within the plasma column. This distinction is by no means to the benefit of traps with magnetic mirrors. One important advantage of adiabatic traps which are open at the two ends is, however, that hot plasma can be produced by a wide variety of methods, of which the best (in principle) is injection of fast particles from outside.

The development of this branch of thermonuclear research has seen some dramatic switches from optimistic expectations to the confounding of hopes. After a long period, during which the experimental data had a disputed and contradictory character, the necessary condition for the stable confinement of hot plasma in traps with magnetic mirrors was established for the first time in 1961. This necessary condition is that the strength of the magnetic field in the trap must increase in all directions away from the region in which the plasma is contained (and essentially this requirement must also be met in closed systems though in a slightly altered form). Experiments with adiabatic traps only began to provide a serious basis for further development after the construction of experimental devices which met this requirement. Such work is at present in progress at Livermore Laboratory in the United States, in the Institute of Atomic Energy in the USSR and also in the French Scientific Centre at Fontenay-aux-Roses. The best plasma parameters so far achieved are $T_i \sim 6 \times 10^7$ and $n\tau_E \sim 7 \times 10^9$. This is a significant step forward when these values are compared with those which characterized the state of this work at the time of the Third Geneva Conference in 1964. It would appear, however, that even under the optimum conditions of comparatively high plasma density ($n_e \sim 10^{13} \text{ cm}^{-3}$), the rate at which it is lost is at least several times larger than the value corresponding to a stable state. Consequently the achievement of the condition $n\tau_E \sim 10^{14}$ in an adiabatic trap may only be possible if an effective method can be developed for recovering the energy lost by the plasma as particles escape through the magnetic mirrors. This energy must be returned into the plasma as a flux of rapid particles from a special injector. In such a closed cycle the efficiency of recirculation has to be higher than 97%; even for a fully stable plasma it would need to be more than 90%. Thus in a thermonuclear reactor working on the principle of adiabatic confinement, the useful power would only be a rather small fraction of the power circulating in the system. Various methods of recovering the energy are under consideration but it is still too early to pass judgment on the feasibility of constructing such a recirculating cycle.

Traps with magnetic mirrors continue to be an important aspect of scientific work on controlled nuclear fusion in the United States and the USSR. It is reasonable to hope that even during the next five years the question of the practicability of such systems will be clarified.

So far I have considered systems for magnetic confinement of a high temperature plasma in a stationary or quasi-stationary state with parameters which change slowly in time. "Quasi-stationary" here signifies that the forces acting on a plasma come to equilibrium, that is the plasma pressure at any instant is compensated by the difference in magnetic pressure. Considerable attention has also been paid to methods of producing hot plasma in strongly non-stationary regimes, where a significant role is played by inertial forces acting on a plasma accelerated by a varying magnetic field. The most highly developed investigations have been on the

so-called "theta pinch", a dense plasma column formed within a long straight tube, in which is created a uniform longitudinal magnetic field, which grows in a few μs to several tens or even hundreds of kgauss. The pulsed magnetic field compresses a cold plasma previously created within the tube and heats it. In such rapid processes it is possible to produce plasma columns with densities of the order of 10^{16} particles cm^{-3} and temperatures of several tens of million K. The confinement time of the energy in the plasma increases with the length of the tube in which it is formed. The largest contributions to the investigation of theta pinches have been the work of the Los Alamos Laboratory in the United States and of the Culham Laboratory in Britain. The best parameters for such systems at present are $T_i \sim 6 \times 10^7$ and $n\tau_E \sim 2 \times 10^{11}$. The basic loss of energy in theta pinches is due to the open ends, and so much attention has been given recently to the possibility of changing from open systems to closed ring-shaped theta pinches. In order to create an equilibrium and conditions for macroscopic stability of the plasma column, the complexity of the structure of the external magnetic field has to be significantly increased. If experience with closed theta pinches proves successful, it opens up the possibility of creating a dense plasma configuration with a comparatively long energy confinement time and a high temperature with an economic use of the magnetic field forces—the ratio of the plasma pressure to the pressure of the magnetic field in the theta pinch should be very much larger than the Tokamak. Macroscopic instability itself is, however, not sufficient for a long confinement of energy in a plasma, for in closed magnetic systems the rate of loss of heat depends on the transverse components of the magnetic field. It is presumably necessary for this component in a closed theta pinch to have a sufficiently large value. It is also necessary that the radius of the perpendicular cross section of the plasma column should not be too small as τ_E is proportional to the square of this radius. Whether these conditions can be met in a toroidal theta pinch is not yet clear, although the high plasma density should play a favourable role. While considering pulsed processes for heating plasma, mention must be made of experiments in which a "plasma focus" is produced, that is a dense blob of hot plasma ($n_e \sim 10^{18}$ to 10^{19} particles cm^{-3} , $T_i \sim 10^7$ K) contained in a very small volume. This arises as a result of the compression of plasma by the magnetic field of a current flowing through it, and it lasts $\sim 0.1 \mu\text{s}$. The plasma focus is the most advanced realization of the concept of magnetic confinement in the direction of processes of very short duration. It can also serve as a pulsed source of neutrons. The quantity $n\tau_E$ in the plasma focus still remains undetermined, however, and it is at present difficult to estimate the practical prospects of using this process.

Laser Heating

Recently it has become clear that there is a possibility of proceeding still further in the direction of extremely rapid processes for generating thermonuclear conditions. In the first place, there is the heating of dense matter by laser pulses. In this method a laser beam is focused on a spot of diameter ~ 0.01 mm on the surface of a thin deuterium target. In most experiments for which data have been published the duration of the light pulse has been less than 10^{-8} s, and its power 10^{13} to 10^{14} W cm^{-2} . On exposure to the light pulse the temperature of the solid target increases rapidly. Initially an extremely dense plasma with a very elevated temperature is formed. The light energy is strongly damped in the relatively thin plasma layer in which resonance conditions are satisfied, that is the frequency of the incident radiation is close to the frequency of natural oscillations of the plasma, which is proportional to $\sqrt{n_e}$. The heat is rapidly transported from this layer to the inside of the target, heating regions with even higher density. Because of the very rapid

conduction of energy within matter, the target can be heated to extremely high temperatures before it disintegrates. It is therefore possible that in such processes magnetic heat insulation is unnecessary. Neutrons resulting from nuclear reactions in matter subjected to a laser beam were first observed in the Physical Institute of the Academy of Sciences in Moscow in 1968; the target in these experiments was lithium deuteride. More recently French physicists at Limeil have managed to raise the yield of neutrons to 10^4 neutrons a pulse. At present the most promising avenue of approach in the development of this method seems to be the heating of a small drop of deuterium (or a mixture of deuterium and tritium) in a uniform field of convergent light generated by nanosecond pulsed lasers. Recently a group of American physicists (E. Teller *et al.*) has proposed a method of obtaining matter in a state of superhigh density by means of a uniformly convergent laser beam. The target is compressed by the reaction force, due to the very rapid evaporation of the target material. Calculations show that with a suitable form of the laser pulse it may be possible to increase the compressive force in the very short final phase of the heating pulse. At very high densities the ignition of a powerful thermonuclear reaction is much more feasible. Investigations of the process of laser heating may soon become one of the principal efforts in research on controlled thermonuclear fusion. The prospects for further development in this direction depend primarily on progress in increasing the efficiency of lasers themselves, which is at present very low. And it is still difficult to predict the laser power which would be required to achieve the values of T_i and $n\tau_E$ which are necessary if the thermonuclear reactions are to give a positive energy yield.

In principle it is possible that extremely rapid heating of matter to thermonuclear temperatures might be achieved by using an extremely powerful and well focused flux of relativistic electrons as the primary energy source. This method of heating might have practical prospects if theoretical predictions of the anomalously rapid deceleration of the electron beam in the plasma are confirmed and if, in addition, difficulties connected with the focusing of powerful electron beams (with currents of up to 10^6 A) can be overcome.

State of Research

In summing up on the state of controlled fusion research, it must first be stated that for many years there has been steady progress along the traditional lines. Naturally, as compared with the initial rosy expectations of the rapid advent of the thermonuclear millenium, this progress may seem very slow, but nevertheless progress there has been. Furthermore, some new approaches to the problem have appeared, and these seem to be highly promising. One way or another, it would seem that there is a reasonable basis for the hope that within the next decade the long sought values of T_i and $n\tau_E$ will be reached, after which the greatest importance will attach to purely engineering problems connected with the construction of the reactor. It is, however, quite unclear which of the several possible solutions to the thermonuclear problem will be the first to reach the goal. He is a brave man indeed who would try to depict the outlines of the future thermonuclear reactor. In the first place, he would have to take a stand on the interval between the present time and the thermonuclear era.

Fusion and Civilization

I have consistently been of the opinion that the problem of controlled nuclear fusion will be solved not later than when it is necessary for the further development of civilization on Earth—if only our children and grandchildren have enough intellect and will to resist the instincts of hostility and destruction.

Man-made Carbon Dioxide and the "Greenhouse" Effect

J. S. SAWYER

Meteorological Office, London Road, Bracknell, Berkshire RG12 2SZ

In spite of the enormous mass of the atmosphere and the very large energies involved in the weather systems which produce our climate, it is being realized that human activities are approaching a scale at which they cannot be completely ignored as possible contributors to climate and climatic change.

The first thing that has to be recognized is that significant effects on the climate are only likely where human activity impinges on a particularly sensitive factor among those that control climate. The output of human industry is still very much less than the total mass of the atmosphere and man-made energy is still small compared with the energy of meteorological systems. The total industrial output of heat each day is, for example, considerably less than 0.1% of the total kinetic energy of the atmosphere, which itself is destroyed by friction and replaced naturally within a few days. Another useful comparison is that of the total man-made heat output in Britain with natural processes over the same area. Even over this area of relatively intense human activity man's efforts are relatively quite small—man-made heat is less than 1% of the energy received from the Sun.

It must also be remembered that the mass of the atmosphere is enormous compared with the products of human activity. The total mass of the atmosphere is more than 500 times the mass of the known coal reserves, for example, and human activities will not significantly change its chief constituents. Nevertheless there are certain minor constituents of the atmosphere which have a particularly significant effect in determining the world climate. They do this by their influence on the transmission of heat through the atmosphere by radiation. Carbon dioxide, water vapour and ozone all play such a role, and the quantities of these substances are not so much greater than the products of human endeavour that the possibilities of man-made influences can be dismissed out of hand.

Influence of CO₂

This article is concerned with the part played by carbon dioxide in determining climate and the way in which it may be affected by human activity. There are several other possible ways in which man might affect climate on a global scale, but the carbon dioxide effect is probably the one about which the most is known, and at the same time it illustrates clearly the inherent difficulties in assessing whether such activity can have a significant effect on climate and how much that effect might be.

First, it is necessary to consider the natural behaviour of carbon dioxide in the atmosphere and the evidence for changes produced by human activity. Carbon dioxide is, of course, a product of the combustion of nearly all fuels and is discharged to the atmosphere through the chimneys or exhaust of the power or heating plant in which the fuel is consumed. The carbon dioxide content of the atmosphere was first measured in the earlier part of the nineteenth century and found to be fairly uniform, both geographically and with season. By the earlier part of the century it had been pointed out that the combustion

of fossil fuel was discharging carbon dioxide, which might be increasing the carbon dioxide content of the atmosphere, and which might have an effect on the heat balance of the Earth. The nineteenth century measurements of carbon dioxide were naturally of a somewhat uncertain accuracy, but by comparison with more recent observations Callendar¹ was able to demonstrate that there is reasonable evidence for an increase attributable to the carbon dioxide added to the atmosphere from the burning of fuel.

The reality of this increase has been confirmed in a remarkable way in the past decade by two series of measurements made at two locations specially chosen to be far away from local sources of pollution. These were at the South Pole and on the summit of the volcanic mountain Mauna Loa in Hawaii. Fig. 1 (taken from ref. 2) shows the trend of carbon dioxide concentration over the past ten years or so at Mauna Loa. The upward trend is apparent, but the diagram also illustrates some other aspects of the problem. The upward trend at the South Pole is closely similar to that at Mauna Loa.

The upward trend amounts to about 0.7 parts per million (p.p.m.) by volume per year over the eleven year period. The value of 312 p.p.m. in 1958 rising to 319 p.p.m. by 1969 compares with the value of around 292 p.p.m. measured by observers in the nineteenth century. Fig. 1 also shows a marked annual fluctuation, emphasizing that carbon dioxide is essential to the growth of vegetation, and is taken up by plants as they grow. (Some is then returned to the atmosphere when they rot and some incorporated in humus, only to be subsequently returned to the atmosphere.) The annual fluctuation in the concentration of carbon dioxide arises because plant growth is greater in the northern hemisphere than in the southern hemisphere (where the land mass is smaller) and plant growth and the uptake of CO₂ is thus a maximum in the northern summer. The carbon dioxide content of the atmosphere is a maximum in the northern spring and a minimum in the northern autumn.

Fig. 1 also shows the rate at which the concentration of CO₂ would have increased if all the man-made carbon dioxide had remained in the atmosphere. The observed increase is only about half of this, the remainder has clearly been removed from the atmosphere by natural processes, and an assessment of the future rise in carbon dioxide content requires a knowledge of what the processes are and where the carbon dioxide is going

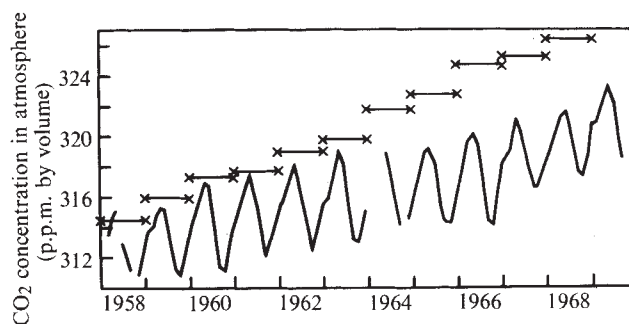


Fig. 1 Increase in carbon dioxide concentration from burning of fossil fuels. —, Mean monthly atmospheric CO₂ concentration at Mauna Loa; x—x, potential annual increase from burning fossil fuels. (Reproduced from ref. 2.)

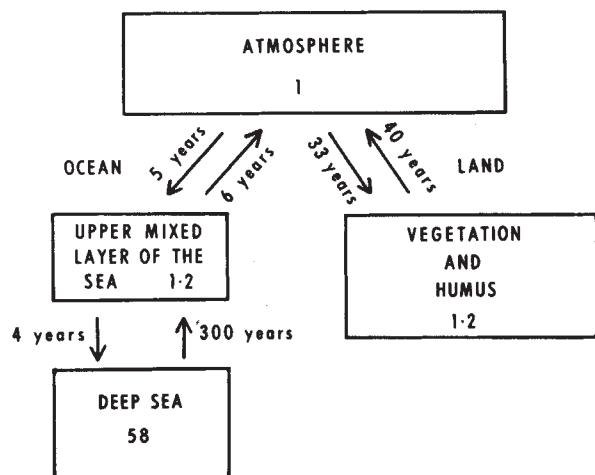


Fig. 2 Natural reservoirs of carbon dioxide (based on ref. 3). Figures indicate content as multiples of atmospheric content.

to. There is also a suggestion in Fig. 1 that the rate of increase of carbon dioxide was rather slower in the mid-1960s than before and after—an indication that these natural processes may vary in their effectiveness from time to time.

Storage and Reservoirs

Fig. 2, which is based on a diagram due to Craig³, shows the natural reservoirs of the carbon which takes part in the carbon dioxide cycle and the relative size of these reservoirs. On land the carbon dioxide is taken up by vegetation and stored in plants and humus. This reservoir is of similar magnitude to that of the atmosphere, and the exchange time is probably of the order of 30 to 40 yr. The ocean provides a much larger reservoir and has the potential for storing some sixty times as much carbon dioxide as the atmosphere. The upper layers of the sea (above the thermocline) must, however, be distinguished from the deeper layers of the ocean. The upper layers are well mixed, and are in contact with the atmosphere, but they can hold only about as much carbon dioxide as exists in the atmosphere. Studies of the concentration of ^{14}C , which is produced by cosmic rays in the atmosphere and subsequently decays to ^{12}C , suggest that the rate of transfer of carbon dioxide from the atmosphere to the upper layers of the ocean is such as to require some 5 to 10 yr for the transfer of a quantity equivalent to that in the atmosphere. Transfer to the deep ocean from the upper layers is a slower process, and as a result it would be a matter of centuries before the deep ocean reached equilibrium with any new level of concentration in the atmosphere.

Industrial development has recently been proceeding at an increasing rate so that the output of man-made carbon dioxide has been increasing more or less exponentially. So long as the carbon dioxide output continues to increase exponentially, it is reasonable to assume that about the same proportion as at present (about half) will remain in the atmosphere and about the same amount will go into the other reservoirs. On this basis Bolin⁴ has estimated that the concentration of carbon dioxide will be about 400 p.p.m. by the year 2000. A recent conference⁵ put the figure somewhat lower (375 p.p.m.)

On the other hand, there must ultimately be a levelling off in industrial output of carbon dioxide, if only on account of the limitations of fuel supply. At this stage a larger proportion of the carbon dioxide will be absorbed by the oceans because, on the longer time scale, the deep sea will have opportunity to come nearer to equilibrium with the atmosphere. If the carbon dioxide were to be shared between the various reservoirs in proportion to their capacities, only one-sixtieth of the man-made carbon

dioxide would remain in the atmosphere—but unfortunately the situation is more complex than that.

Chemical Complications

Kanwisher⁶ has pointed out that only a small proportion of the carbon dioxide entering the sea remains as dissolved CO_2 directly available for exchange with the atmosphere. The remainder forms magnesium and sodium carbonates which provide a chemical buffering solution for carbon dioxide. In consequence an increase of 0.6% in the carbon dioxide content of the sea corresponds to a 10% increase in the partial pressure of CO_2 in the atmosphere above, and on this basis one might expect that the ultimate sharing of carbon dioxide between the atmosphere and the whole ocean might still leave more than 20% of the extra carbon dioxide in the atmosphere. On a still longer time scale one might expect some of the oceanic carbon to be deposited as carbonates on the sea bed, but the typical times associated with this process is probably too long to be of much relevance to the fate of industrially produced CO_2 .

There is little doubt that in assessing the future level of carbon dioxide in the atmosphere, it is important to understand fully the balance between the carbon dioxide in the atmosphere and in the ocean. There are several other complications which are not fully understood. The solubility of carbon dioxide is greater at lower temperatures and the tropical oceans are thus continuously giving up carbon dioxide to the atmosphere; it is then reabsorbed in the oceans of higher latitudes. In most parts of the ocean there is a relatively warm layer of water lying above colder and denser water beneath—the transition layer, known as the thermocline, lies at a depth of 100 to 200 m. This stable layer is a barrier to the mixing of the lower and upper water of the ocean, but the barrier disappears in certain parts of the polar ocean when the surface water is cooled in winter, and such areas may provide a pathway by which carbon dioxide absorbed from the atmosphere can be transferred to the deep ocean more readily. Such areas may play a significant role in determining the carbon dioxide balance between ocean and atmosphere, and an understanding of this balance will require a better understanding of the long term circulation of the ocean than is available at present.

Indirect Effects of Increased CO_2

The direct effect of a small increase in carbon dioxide on mankind would be negligible (except that it might make some vegetation grow a little faster) and I shall now consider its possible indirect effect on the world climate.

The temperature of the Earth is, of course, maintained by the energy received from the Sun as radiation in a band of wavelengths centred in the visible region. Some of this radiation is reflected by the Earth's surface, and more especially by clouds, but most of the remainder penetrates the atmosphere and heats the Earth's surface including the ocean. Some of the heat is radiated back by the surface at longer wavelengths corresponding to its lower temperature than the Sun, part is communicated to the overlying air by conduction, and part is used to evaporate water and becomes available to heat the air when the water condenses as rain. The atmosphere is not transparent to the long wavelength radiation which is emitted by the Earth and its atmosphere, as it is to the incoming short wave radiation. Thus certain atmospheric gases, principally water vapour and carbon dioxide, absorb a significant part of the outgoing radiation and reradiate it both upwards and downwards. (The significant aspects of the radiation spectrum are illustrated in Fig. 3.) The outgoing radiation from the Earth-atmosphere system is thus made up of, first, some radiation emitted by the Earth's surface at wavelengths to which the Earth's atmosphere is transparent—the so-called "window", primarily between 7 and 14 μm ; second, some radiation which has been emitted from the surface (or from clouds), absorbed by atmospheric gases and reradiated outward by the same (or other) gases; and, third, radiation from clouds

which themselves may be receiving heat from below. Some of the heat radiated outward by the gases and clouds is transported from below to the level where it is radiated. These aspects of the radiation balance of the atmosphere are illustrated in Fig. 4.

As carbon dioxide is one of the principal gases taking part in radiation exchange in the atmosphere and in the radiation of the Earth's heat content, a change in the content of carbon dioxide within the atmosphere is likely to influence the process. The chief effect of increasing carbon dioxide is that the gas which is radiating heat to space is found at a higher level in the atmosphere than before—the radiation from lower down in the atmosphere is absorbed by the extra carbon dioxide above and then reradiated to space. In the troposphere, at least, temperature decreases with height, so the effective radiating temperature of the carbon dioxide becomes lower if the amount of the gas is increased, and therefore less heat is radiated to space. Thus the additional carbon dioxide tends to act as a blanket which keeps the Earth warmer—the Earth has to get rid of the incoming radiation from the Sun, and the same amount can only be removed if the temperature of the atmosphere rises a little.

The calculation of the effect on the radiation budget is not easy because the absorption and emission of heat through the whole of the long and short wave radiation spectrum has to be taken into account, and the heat transfer has to be calculated by integration over the very complicated spectrum resulting from the complex pattern of absorption bands introduced by the various atmospheric constituents. Nevertheless modern computing resources allow reasonably accurate computations to be made. Probably the most reliable calculations performed so far are those reported by Manabe and Wetherald⁷, who point out and overcome two important deficiencies of earlier calculations. The deficiencies arise, first, because the vertical distribution of temperature in the atmosphere is not determined by radiation transfer alone; and, second, because, if world temperature rises due to an increase in carbon dioxide, it is almost certain that there will be more evaporation of water—the water vapour content of the atmosphere will also increase and will have its own effect on the radiation balance.

The first difficulty is overcome by incorporating a restriction in the calculations, namely, that the temperature must not fall off more rapidly with height than $-6.5^{\circ}\text{C km}^{-1}$, the rate normally found in the troposphere. If this rate is exceeded during the calculations a so-called "convective adjustment" is made.

The second effect is probably more important in the present context. An atmosphere at a higher temperature can hold more water vapour, and the additional water vapour produces a similar blanketing effect to that produced by carbon dioxide. Manabe and Wetherald⁷ calculate that an increase of 100% in

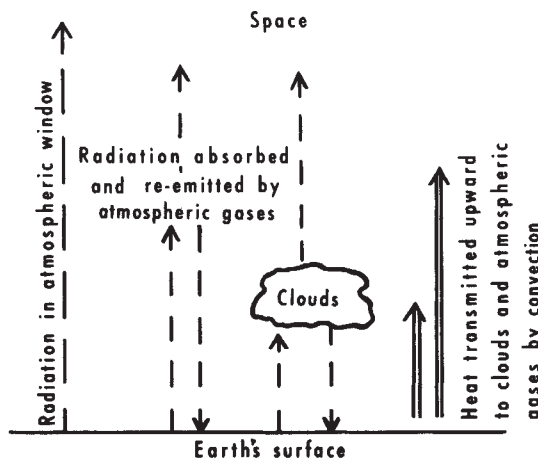


Fig. 4 Pathways of heat loss from the Earth and atmosphere.

the content of carbon dioxide would increase the world temperature by 1.3°C if the water content of the atmosphere remained constant, but by 2.4°C if the water vapour increased to retain the same relative humidity. The increase of 25% in CO_2 expected by the end of the century therefore corresponds to an increase of 0.6°C in world temperature—an amount somewhat greater than the climatic variations of recent centuries. Rasool and Schneider⁸ give a more recent and substantially smaller estimate of the effect of carbon dioxide, but their scheme makes no allowance for adjustments of stratospheric temperature or absorption of incoming radiation by carbon dioxide.

Other Variables

The water vapour content is by no means the only variable factor in the atmosphere which would change as a result of a general warming. The increased water vapour would probably lead to the formation of more clouds because evaporation increases much faster than temperature, and substantially more condensed water would be available. The additional cloud would reflect incoming solar radiation and tend to produce a lowering of temperature—a negative feedback effect to counterbalance the positive feedback arising from water vapour. Other calculations⁷ show that world temperature is likely to be remarkably sensitive to the average global cloudiness. A change of only 1% in the average cloudiness would produce a change in temperature of almost 1°C . As cloudiness varies greatly from day to day and from place to place, it is really rather remarkable that the fluctuations in world temperature are not greater than those observed. There is thus a suggestion that the negative feedback from cloud cover is a real stabilizing effect on world temperature, increasing cloud keeping the temperature down and cutting off evaporation until the average cloudiness is restored. No quantitative estimate of this effect has, however, yet been made.

On the other hand, a destabilizing feedback which has been given some prominence recently arises from the changes in the Earth's reflectivity which would accompany any change in its permanent cover of ice and snow. A lowering of temperature would lead to an extension of glaciers and permanent snow fields, and these would reflect more of the incoming solar radiation. Budyko⁹ has stressed the destabilizing influence of such positive feedback and has suggested that a decrease of incoming solar radiation of only about 2% might lead to another ice age and indeed to complete glaciation of the Earth. The effect of the decrease of ice cover which would accompany an increase of world temperature is, however, less significant. The present limits of permanent snow and ice are such that no extensive recession is possible with increasing temperature without a thawing of the ice cover of the Arctic Ocean. If this took place it would undoubtedly have a significant effect on the

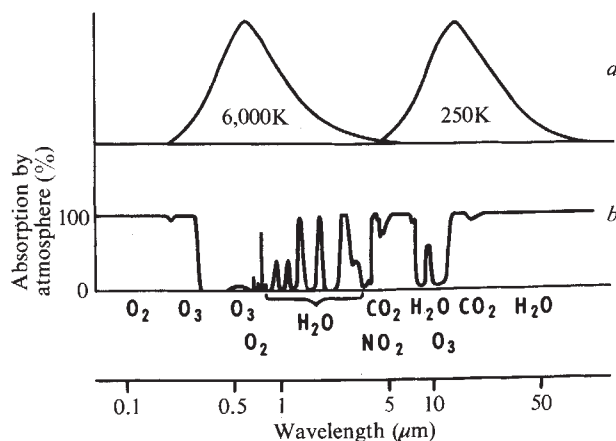


Fig. 3 Radiation spectrum illustrating absorption by principal atmospheric absorbing gases. a, Black body curves; b, absorption by atmosphere.

climate of the adjacent coasts and hinterland, but the indications are that the effect on the mean temperature in other latitudes would be limited to 1° or 2° C. Regional climatic effects might, however be significant.

The oceans also have an influence on climatic changes through their large mass and thermal inertia. The atmosphere cannot settle down to a new temperature regime until the temperature of the oceans has come into equilibrium with that regime. Calculations show that this would take of the order of 100 yr, and in consequence the oceans impose a substantial lag on the response of world temperature to such changes as I have discussed here. On the other hand, a rise in temperature of the oceans will itself release additional carbon dioxide to the atmosphere and provide a positive feedback tending to enhance slightly the carbon dioxide effect.

The response of the atmosphere to a change in its heat balance is far from simple. Almost any change in the heat supply is likely to result not only in a change in the overall temperature, but also in the system of winds and weather which derive their energy from the heating and cooling of the atmosphere. Such a change in the winds produces of itself a change in climate, and also modifies the cloud and temperature distribution with a possible feedback to the heat supply and heat loss by radiation.

In spite of the enormous complications of attempting to calculate the whole circulation of atmospheric winds and the resulting cloud and rain distributions, it is probable that only in this way can a soundly based estimate of possible man-made climatic changes be made. Numerical models of the atmospheric circulation have been developed and the computations are just practicable on the largest modern computers. The calculations reproduce the chief features of the world climate, but do not have sufficient precision to differentiate two regimes which would produce a difference of temperature of 1° or 2° C. Possibly such calculations will be feasible in the future, but the sophisticated models required will have to take into account the complicated feedback processes which I have already discussed. Most difficult is to devise a method of calculating the amount of cloud to be expected in a particular circulation regime because individual clouds are too small to be treated; the statistical behaviour of assemblies of clouds covering a region will have to be calculated, and it is not yet clear how this may be achieved.

Possible Changes and Natural Fluctuations

The climate has undergone many changes in the past. Some of these have been associated with the formation of the atmosphere itself, the generation of oxygen by vegetation and so on,

but large scale geographical changes, the moving of the continents, have also played a part. Even since geography and atmospheric composition have been more or less as they are known at present, there have been ice ages and periods considerably milder than the present—the causes of these changes are unknown. The last ice age came to an end over a period of centuries about 10,000 yr ago, but even since then there have been changes. Since the temperature recovered to more or less current values the climatic fluctuation in England has at most been a matter of 2° C or so. The biggest swing in recent decades has probably been about 1° C from the medieval optimum to the little ice age around the seventeenth century. Individual years, even in the past two decades or so, have, however, covered a wider range than 2° C. Even global mean temperatures have varied by 0.6° C from a minimum around 1880 to the last maximum around 1940. Against this background a change of 0.6° C by the end of the century will not be easy to distinguish from natural fluctuations and certainly is not a cause for alarm. Even a doubling of the amount of carbon dioxide in the atmosphere, which would probably require the burning of a large part of the known fuel reserves, would appear to result in a rise of temperature little above that experienced in the climatic optimum which followed the last ice age. Nevertheless it must not be overlooked that variations in climate of only a fraction of a degree centigrade have considerable economic importance, as experience of natural fluctuations has already shown. The more frequent incidence of severe winters or of frost can readily affect the economies of sensitive crops.

Although there may be no immediate cause for alarm about the consequences of carbon dioxide increase in the atmosphere, there is certainly need for further study. We need to have a better assessment of where the carbon dioxide goes after it has been dispersed from our chimneys, and in particular the long term balance with the ocean. We also need a better assessment of the various feedback effects which enter into the control of climate, and the development of increasingly sophisticated numerical simulation of the global climate seems the only possible approach in spite of the effort involved.

- ¹ Callendar, G. S., *Quart. J. Roy. Met. Soc.*, **66**, 395 (1940).
- ² *Man's Impact on the Global Environment* (MIT Press, Cambridge, Massachusetts, 1970).
- ³ Craig, H., *Tellus*, **9**, 1 (1957).
- ⁴ Bolin, B., *Stockholm, Stat. Naturvetensk. Forskningsgr., Sartryck. Svensk Naturevetensk.*, 134 (1966).
- ⁵ *Inadvertent Climate Modification* (MIT Press, Cambridge, Massachusetts, 1971).
- ⁶ Kanwisher, J., *Tellus*, **12**, 209 (1960).
- ⁷ Manabe, S., and Wetherald, R. T., *J. Atmos. Sci.*, **24**, 241 (1967).
- ⁸ Rasool, I., and Schneider, S. H., *Science*, **173**, 138 (1971).
- ⁹ Budyko, M. I., *Tellus*, **21**, 611 (1969).

Progress Maintained in Molecular Biology

P. W. J. RIGBY

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH

Although progress in molecular biology may be slowing down somewhat, an account of recent developments inevitably involves some selection.

SUFFICIENT important developments have occurred in molecular biology during the past year to make it difficult to review the

subject comprehensively in an article such as this. The selection that has occurred is bound to reflect personal interest, but the topics discussed here probably comprise the more important achievements of the past year, together with those lines of experimentation that seem most likely to lead to interesting results.

The determination of the complete nucleotide sequence of a structural gene represents a most impressive technical achievement by Fiers and his group in Ghent; they have reported the

sequence of the coat protein cistron of the RNA coliphage MS 2 (ref. 1). Moreover, their sequence information leads to several conclusions of fundamental significance. The sequence, which was determined chiefly by the techniques developed by Sanger and his colleagues, covers the untranslated region preceding the coat protein cistron, the whole of that cistron, a second untranslated region, and then the beginning of the RNA replicase cistron. The sequence can be arranged into a secondary structure, called the "flower" model because it resembles a stalk with petals, in which 65% of the bases are paired. As with R17, the initiating AUG for the coat protein is found at the top of a base paired loop, but that for the replicase is not; sequence evidence has not yet provided any coherent picture of how ribosomes pick out the correct AUG at which to initiate polypeptide synthesis. The most striking success of the proposed structure is that it provides an explanation for the polarity of certain amber mutants. Amber mutations, which cause premature termination of protein synthesis, are polar at position 6 of the coat protein cistron—that is, they do not allow translation of the subsequent RNA replicase cistron. Ambers at later positions—for example, 50—are, however, not polar. This is explained by the fact that the initiating AUG for the replicase is in a base paired region involving the RNA that codes for amino-acids 24 to 32 of the coat protein. Thus it can be postulated that translation of this region is required to break up the base pairing and thus expose the initiator codon. These results serve to confirm the idea that messenger structure has an important role in the control of translation, but the generality of the concept cannot be assessed until the sequence of an intracellular messenger is elucidated—this is a problem that is being attacked in several laboratories.

Sanger's group at the MRC Laboratory of Molecular Biology in Cambridge have continued their studies on the closely related phage R17 and have recently described further sequences from its coat protein cistron². This group has now turned their attention to the problem of DNA sequencing. The successful development of a methodology for the total determination of DNA sequences would allow investigation of many pressing problems in molecular biology, particularly those concerned with the nature of control signals encoded in DNA. The results published so far³ are concerned with the determination of limited sequences, those of pyrimidine tracts from coliphages. Such a limited approach is necessary because of the absence at present of reproducible procedures for the specific enzymatic digestion of DNA.

The approach used involves the chemical depurination of ³²P-labelled DNA and then fractionation of the resultant oligonucleotides, using the techniques developed for RNA sequencing. These oligonucleotides are then subjected to limited digestion, using exonucleases that attack specifically from one end or the other of the DNA. The digestion products are then fingerprinted and, as they contain a complete variety of products differing by one residue, ranging from undegraded starting material to mononucleotides, and as the fingerprinting procedure separates on the basis of size and base composition, the complete sequence can be read directly from the fingerprint. The largest oligonucleotide sequenced is 20 long, but unfortunately it can be shown not to be part of the cistron coding for the coat protein, the one protein of phage fd of which the amino-acid sequence is known.

The techniques of nucleotide sequencing developed in Cambridge have been successfully applied to an investigation of the mechanism of biosynthesis of tRNA. This work represents a most elegant exploitation of the tyrosine suppressor system, which has been extensively studied in Cambridge over the past few years by Smith and his colleagues. Smith and Altman reported the isolation of mutants of *Escherichia coli* which accumulate a molecule that is a precursor to mature tRNA (ref. 4). It has been known for some time that ribosomal RNA in both bacteria and mammals is transcribed from the gene in units longer than the final functional product. Large quantities of this precursor can be obtained by infecting

cells with a derivative of phage ϕ 80 carrying a mutant tyrosine tRNA gene. The molecule contains the complete nucleotide sequence of the mature tRNA together with an additional 41 bases at the 5' end and three additional bases at the 3' end. It is completely unmodified and, as the 5' terminal residue is pppGp, is likely to represent the initial transcription product. It is interesting to note that the precursor does contain the CCA_{OH} trinucleotide that is at the 3' terminus of the mature tRNA. This turns over within the cell and it had often been suggested that it was not a primary transcription product.

A further mutant that accumulates more precursor was analysed and was shown to carry a base change in a position very close to the cleavage point. The level of precursor is affected by mutations both close to and remote from the cleavage point and a scheme for tRNA maturation is proposed in which both base sequences and secondary structure play a part. A brief characterization has been performed of a ribosome-associated endonuclease which appears to cleave off the extra nucleotides at the 5' end but not those at the 3' end. Further studies on this enzyme and on the modification and maturation of the tRNA will be awaited with great interest.

Protein Biosynthesis

The enzymology of protein biosynthesis has continued to attract considerable attention and work from a number of laboratories, which has led towards the conclusion that there is only one GTPase moiety on the ribosome that catalyses the GTP hydrolysis associated with a number of synthetic steps. This concept was originally proposed by Modolell and his colleagues⁵ as a result of *in vitro* studies which showed that the antibiotic thiostrepton inhibits both the formation of ribosome-EFG-GDP complexes and the GTPase activity associated with EFTu directed binding of aminoacyl-tRNA to ribosomes, and by Cundliffe⁶ as a result of *in vivo* studies on the mode of action of thiostrepton. The *in vivo* results also suggest that the GTPase is located in the acceptor site of the ribosome.

These results stimulated a large amount of research and several groups have confirmed and extended them. Richter⁷ showed that when EFG and GDP are bound to 70S ribosomes in the presence of fusidic acid they cannot accept the ternary complex of EFTu-GTP-Phe-tRNA. Fusidic acid stabilizes the ribosome-EFG-GDP complex; *in vivo* studies⁸ indicate that it, like thiostrepton, inhibits the functions of both EFG and EFTu. If the abortive EFTu-GMP-PCP-Phe-tRNA complex is bound first then the capacity of the ribosomes to bind EFG and GTP is greatly reduced. Richman and Bodley⁹ showed that if the EFTu-GTP-tRNA complex is bound to 30S subunits, a process that does not involve GTP hydrolysis, then EFG, GDP and fusidic acid have no effect.

Grunberg-Manago and her colleagues have extended these observations to the GTPase activity associated with initiation¹⁰. The IF2 directed binding of f-met-tRNA_f to ribosomes is normally concomitant with the hydrolysis of GTP; this hydrolysis is inhibited by thiostrepton although it had previously been shown by Modolell's group that the antibiotic does not inhibit the binding of the initiator tRNA. These results suggest that initiation does indeed involve the binding of f-met-tRNA_f into the peptidyl site of the ribosome. It would thus seem that there is one GTPase activity on the ribosome that can interact with factors IF2, EFG and EFTu. As yet there is no evidence for the involvement of GTP in chain termination in bacteria, although GTPase activity is probably associated with the process in mammals¹¹.

Preliminary results are available concerning the location on the ribosome of this GTPase activity. Washing ribosomes in high salt and ethanol removes several proteins, including one that is required for EFG-dependent GTP hydrolysis¹². This protein, which is located on the 50S subunit, is extremely acidic and has a molecular weight of 13,000 daltons. It is also required for EFTu-dependent GTPase activity. Studies with

antibiotic binding indicate that this protein may be sufficient only for the binding of the transfer factors and that another protein is required for enzymatic activity. A preliminary protein chemical characterization has been reported¹³. Further studies on the structure and function of this A protein will be awaited with great interest. Results available at present indicate that it may take a central role in the action of the ribosome, particularly in the hitherto mysterious process of translocation.

The replication of the RNA of coliphages such as Q β has long been a puzzle, as the purified replicase comprises four polypeptide chains, three of which are coded by the host bacterium. In an elegant study by Weber's group at Harvard it has been shown that subunits III and IV of the replicase are identical to the protein synthesis factors EFTu and EFTs (ref. 14). The identity of the replicase components with the elongation factors has been shown by electrophoresis, by immunological cross-reaction, by functional cross-reactivity, and in one case by amino-acid sequence data. It is also shown that subunits III and IV bind ppGpp, the so-called "Magic Spot I", which is involved in the regulation of RNA synthesis in stringent bacteria. It had previously been suggested by Travers and his colleagues that subunits III and IV, known as ψ factor, are necessary for RNA polymerase to synthesize ribosomal RNA, a process that is inhibited by ppGpp. This idea has since been challenged¹⁵ but the involvement of EFTu and EFTs in both polypeptide synthesis and RNA phage replication, and possibly in host RNA synthesis, suggests that protein and RNA synthesis may well be more closely connected than was previously thought, and the link is likely to be ppGpp. This idea is further supported by Revel's report that subunit I of the replicase is identical to protein synthesis initiation factor i, which is supposed to play a part in the selection of initiation points.

Cyclic AMP

The year has also seen considerable progress in the detailed elucidation of the way in which cAMP controls the expression of inducible genes in *E. coli*. Workers at the National Institutes of Health described¹⁶ an *in vitro* transcription system which is programmed by the DNA of a λ transducing phage carrying the lac operon. Synthesis of lac mRNA on this template requires the presence of RNA polymerase containing sigma, rho factor, crp, the cAMP receptor protein and cAMP. In the absence of either cAMP or crp no lac mRNA is produced. Transcription is blocked by the presence of the lac repressor and this block is alleviated by the addition of inducers such as IPTG. Evidence is presented that indicates that cAMP and crp are required for the binding of RNA polymerase to the lac promoter; crp will bind directly to lac DNA. The crp protein has now been purified to homogeneity and several of its properties investigated¹⁷: it is a dimer of molecular weight 45,000 daltons, comprising two probably identical subunits. The results described above for the lac operon have been essentially repeated using DNA from a λ derivative carrying the gal operon.

Another interesting role of this ubiquitous molecule is in the regulation of the decision between lysogeny and lysis for a temperate bacteriophage¹⁸. It has been shown that mutants of *Salmonella typhimurium* defective in the gene for adenyl cyclase or in the crp gene are lysogenized at a much reduced frequency by phage P22. If exogenous cAMP is added to the cyclase deficient mutants then they are lysogenized at the normal frequency. Some of the genes which must be active for lysogeny to occur are presumably cAMP dependent. This mechanism can be well accounted for in an evolutionary sense, for when the bacteria are poorly supplied with nutrients the internal cAMP concentration will be high and so phage will tend to lysogenize, whereas in conditions of good nutrient supply, when the death of cells is not so critical, the cAMP level will be low and lytic infection will proceed. Ames and his

colleagues also report that mutants resistant to rifampicin, and more surprisingly mutants requiring pyridoxine, are lysogenized at reduced frequencies.

Tumour Viruses

Progress in tumour virology has continued to be rapid. It had been supposed that transformation by oncogenic viruses was in some way a result of the stable integration of the viral genome into the chromosomes of the host, and the continued expression of at least a part of that genome, but it is now clear that viral integration occurs during the lytic infection of several cell lines by polyoma virus^{19,20}. Thus integration *per se* is not sufficient for transformation. It has been shown, moreover, that during lytic infection of monkey kidney cells by SV 40 the virions produced contain cellular DNA covalently incorporated into the viral genome²¹. This incorporation does not occur if plaque purified virus is used, and it is dependent on the multiplicity of infection. Electron microscopic studies²² indicate that the cellular sequences are incorporated at the expense of viral sequences; this explains why the virions produced are defective in plaque formation.

It is obviously a necessary condition for transformation to occur that only certain of the viral genes can be expressed, and thus the control of transcription is likely to be important in the transformation process. Two groups have demonstrated, using nucleic acid hybridization techniques, that *in vivo* transcription of the SV 40 genome is asymmetric^{23,24}. It differs, however, from *in vitro* transcription using the *E. coli* polymerase in that both strands are transcribed. Early *in vivo* RNA is synthesized from the same strand as that read by the bacterial polymerase, whereas late *in vivo* RNA is made on the opposite strand. The demonstration that strand selection occurs in tumour viruses adds a further set of control operations to those that could be involved in transformation.

- ¹ Min Jou, W., Haegeman, G., Ysebaert, M., and Fiers, W., *Nature*, **237**, 82 (1972).
- ² Jeppesen, P. G. N., Barrell, B. G., Sanger, F., and Coulson, A. R., *Biochem. J.*, **128**, 993 (1972).
- ³ Ling, V., *J. Mol. Biol.*, **64**, 87 (1972).
- ⁴ Altman, S., and Smith, J. D., *Nature New Biology*, **233**, 35 (1971).
- ⁵ Modolell, J., Cabrer, B., Parmeggiani, A., and Vazquez, D., *Proc. US Nat. Acad. Sci.*, **68**, 1796 (1971).
- ⁶ Cundliffe, E., *Biochem. Biophys. Res. Comm.*, **44**, 912 (1971).
- ⁷ Richter, D., *Biochem. Biophys. Res. Comm.*, **46**, 1850 (1972).
- ⁸ Cundliffe, E., *Biochem. Biophys. Res. Comm.*, **46**, 1794 (1972).
- ⁹ Richman, N., and Bodley, J. W., *Proc. US Nat. Acad. Sci.*, **69**, 686 (1972).
- ¹⁰ Grunberg-Manago, M., Dondon, J., and Graff, M., *FEBS Lett.*, **22**, 217 (1972).
- ¹¹ Beaudet, A. L., and Caskey, C. T., *Proc. US Nat. Acad. Sci.*, **68**, 619 (1972).
- ¹² Brot, N., Yamasaki, E., Redfield, B., and Weissbach, H., *Arch. Biochem. Biophys.*, **148**, (1972).
- ¹³ Terhorst, C., Wittmann-Liebold, B., and Möller, W., *Eur. J. Biochem.*, **25**, 13 (1972).
- ¹⁴ Blumenthal, T., Landers, T. A., and Weber, K., *Proc. US Nat. Acad. Sci.*, **69**, 1313 (1972).
- ¹⁵ Haseltine, W. A., *Nature*, **235**, 329 (1972).
- ¹⁶ de Crombrughe, B., Chen, B., Anderson, W., Nissley, P., Gottesman, M., Pastan, I., and Perlman, R. L., *Nature New Biology*, **231**, 139 (1971).
- ¹⁷ Anderson, W. B., Schneider, A. B., Emmer, M., Perlman, R. L., and Pastan, I., *J. Biol. Chem.*, **246**, 5929 (1971).
- ¹⁸ Hong, J.-S., Smith, G. R., and Ames, B. N., *Proc. US Nat. Acad. Sci.*, **68**, 2258 (1971).
- ¹⁹ Babiuk, L. A., and Hudson, J. B., *Biochem. Biophys. Res. Comm.*, **47**, 111 (1972).
- ²⁰ Ralph, R. K., and Colter, J. S., *Virology*, **48**, 49 (1972).
- ²¹ Lavi, S., and Winocour, E., *J. Virol.*, **9**, 309 (1972).
- ²² Tai, H. T., Smith, C. A., Sharp, P. A., and Vinograd, J., *J. Virol.*, **9**, 317 (1972).
- ²³ Lindstrom, D. M., and Dulbecco, R., *Proc. US Nat. Acad. Sci.*, **69**, 1517 (1972).
- ²⁴ Khoury, G., and Martin, M. A., *Nature New Biology*, **238**, 4 (1972).

Excitable Membranes

RICHARD D. KEYNES

Agricultural Research Council Institute of Animal Physiology, Babraham, Cambridge

A fluid mosaic model for membrane structure well accommodates the ideas about the mechanism of excitability which arose from the electrophysiological experiments of Hodgkin and Huxley. Recent experiments have been aimed at clarifying the molecular basis for this mechanism.

It is now generally accepted that the conduction of impulses in nerve and muscle fibres depends on their cable structure and on the occurrence of specific changes in the ionic permeability of the surface membrane. The first description of nerve excitation in terms of a sequence of time and voltage-dependent ionic conductance changes was given in the classical papers of Hodgkin and Huxley¹ in the early 1950s. This has stood the test of time with impressive success, and the subsequent work, both theoretical and experimental, in other laboratories has served more to fill out the details of their picture than to provide alternative hypotheses.

Structure of Nerve Membranes

The equivalent electrical circuit of the membrane consists of a capacity in parallel with batteries representing the Nernst equilibrium potentials E_{Na} and E_K for sodium and potassium, each in series with a variable resistance. An additional battery and a fixed "leak" resistance must also be included to cover the contribution to the membrane current from chloride and other unspecified ions. This is relatively small in the squid giant axon, but in tissues like frog muscle the chloride conductance may be greater than the resting potassium conductance². In what may be regarded as the experiments that initiated the whole modern development of electrophysiology, Cole and Curtis³ showed that in a squid giant axon inserted in one arm of an impedance bridge, there was no measurable change in the capacitive element during the passage of an impulse, but a large one in the resistive element. This important observation establishes at once the principle that the components of the membrane basically responsible for its impermeability to small water-soluble molecules and ions do not participate directly in the conduction mechanism, and leads me to consider what the membrane structure is now thought to be.

As far as this impermeable component is concerned, the model generally attributed to Davson and Danielli still holds the field, the proposition being that the matrix of the membrane is composed of a bilayer of phospholipids and cholesterol, arranged with the polar and therefore hydrophilic headgroups facing outwards. It is presumably the headgroup region that is stained with osmium tetroxide to produce the characteristic tramlines spaced about 25 Å apart seen in electron micrographs, for⁴ this is roughly the thickness of the hydrophobic layer that would correspond

to a dielectric constant of 2 and the observed capacity of $1 \mu F cm^{-2}$. The proteins of the membrane have often been depicted in continuous monolayers on either side, forming a trilaminar sandwich. But Singer^{5,6} has given good reasons for preferring a fluid mosaic model, in which the proteins have hydrophobic portions loosely associated with the hydrocarbon chains of the phospholipids, and hydrophilic portions projecting into the aqueous phase. In some cells, the surface antigens have been convincingly shown by Frye and Edidin⁷ to be free to diffuse laterally in the membrane, and are probably arranged asymmetrically so that they are accessible only from the outside. Those proteins and enzyme systems involved in the selective transfer of ions and other substances across the membrane seem likely to have a dumb-bell form, with two hydrophilic ends linked by a central section which presents lipophilic side chains to the phospholipids but has a hydrophilic core. They may be more firmly anchored; there is no evidence one way or the other. They have a variety of special roles in connexion with transmembrane transport, and I am concerned here only with those that constitute the pores or channels through which sodium and potassium ions move downhill during the nerve impulse.

As has been emphasized by Gitler⁸ and others, the fluidity of the membrane is clearly an important concept. The introduction of methods for labelling different parts of it either with paramagnetic⁹ or with fluorescent¹⁰ groups, and nuclear magnetic resonance studies of natural and artificial membranes¹¹, have shown that there is greater freedom of movement near the ends of the hydrocarbon chains in the central region than there is in the neighbourhood of the headgroups of the phospholipids. The liquidity at the centre seems to be appreciably greater in excitable membranes than in myelin or the membranes of cells like erythrocytes. From this fact, and from the effects of changing the proportion of cholesterol or adding local anaesthetics to phospholipid preparations, it seems probable that excitability depends on an environment fluid enough to allow the conformational changes at the active sites to take place without mechanical hindrance. The price paid for this flexibility is greater leakiness; but this can be offset by a greater expenditure of energy on pumping mechanisms. I do not mean to suggest, however, that nerve membranes must be considered to be sufficiently liquid for mobile carrier molecules, rather than some kind of selective pore, to transport the ions. Lowering the temperature of nerve membranes does not result in the abrupt reduction in the ionic conductance which has been described for artificial bilayers incorporating nonactin¹². The conclusion^{13,14} that the entropy of the membrane in rabbit vagus nerve decreases when the potential across it is reduced raises a difficulty in relation to its liquidity, because this is a change in the opposite direction to that which would obtain if the dielectric behaved like a simple liquid. But alterations in membrane thickness also have to be taken into account, and presumably the contribution to the total entropy of the system of an increased disorder of the hydrocarbon chains of the phospholipids when the electric field is reduced is more than counterbalanced by an increased order elsewhere in the structure.

Number of Ionic Channels

An essential feature of the conduction mechanism is that the sodium permeability first rises and then falls again, as in its turn the potassium permeability increases. It was therefore always an attractive possibility that there might be a single set of permeability channels the selectivity of which was modified sequentially, so that after opening with a high sodium conductance they would pass to a state of high potassium conductance before finally closing. But Hodgkin and Huxley¹ soon decided that this idea would be hard to reconcile with the precise timing of the conductance changes. There is now experimental evidence of several kinds that seems to establish beyond reasonable doubt that the sodium and potassium channels are quite separate entities. The arguments have been well summarized by Hille¹⁵, the most telling point being that the sodium conductance mechanism can be blocked completely by very low concentrations of tetrodotoxin without any perceptible change in the behaviour of the potassium system, while conversely the potassium channels can be put out of action by quaternary ammonium ions without affecting the sodium system.

The high specificity of tetrodotoxin, and the remarkably low concentration at which it acts—the equilibrium dissociation constant being around 3×10^{-9} M—can be exploited to measure the binding capacity of nerve membranes for the toxin, and hence to set an upper limit to the density of packing of the sodium sites. This was first done by Moore, Narahashi and Shaw¹⁶, and the measurements have since been repeated with greater accuracy by Ritchie, Rojas and myself¹⁷. Our results suggest that in $1 \mu\text{m}^2$ of nerve membrane the number of sodium sites is not more than seventy-five for rabbit vagus, forty-nine for crab leg nerve, and thirty-six for lobster nerve. Another experimental approach, which involves determination of the slowing down of the escape of tetrodotoxin from the membrane when the binding sites act as a saturable sink and the Schwann cell limits the rate of diffusion, seems to give rather similar numbers for the squid giant axon (unpublished results of R. D. K., E. Rojas, and R. E. Taylor), but the analysis of our results is still far from complete. These figures may eventually have to be revised downwards, because although experiments with tritiated tetrodotoxin give good agreement for the total binding in rabbit nerve¹⁸, they also indicate that part of it is unspecific and may not represent attachment to sodium sites. Moreover, it seems from studies of the kinetics of the uptake and release of tetrodotoxin at the node of Ranvier in myelinated nerve fibres¹⁹ that the slow recovery observed in the squid giant axon²⁰ is only partly due to a diffusion effect.

Within a $1\text{-}\mu\text{m}^2$ area there are about 2×10^6 phospholipid molecules at each face of the membrane, so these estimates of the number of sodium channels are strikingly small. It can be calculated that the sodium conductance of a single channel is around 1×10^{-10} ohm⁻¹, a value which also applies to the acetylcholine receptor channels at the motor end-plate²¹ and to the unit conductance channel of alamethicin²².

This means that the rate of passage of sodium ions through an open channel is of the order of 10^8 s⁻¹. The maximum rate at which ions could arrive by diffusion at the mouth of the channel is only about ten times greater. Thus the sodium channels should be regarded as conducting pores rather than mobile carriers shuttling to and fro across the membrane, for a carrier mechanism would probably operate much more slowly²³. It would be of interest in this connexion to obtain comparable information about the potassium channels in nerve membranes. But although there are indications that they may be more numerous, and hence may have a lower individual conductance¹⁵, direct counts have not yet been made.

Optically Detectable Structural Changes

Three different optical techniques have recently been used to look for structural changes associated with electrical activity in excitable tissues. Thus measurements have been made of changes in birefringence²⁴⁻²⁷, light scattering^{24,28,29}, and the fluorescence emission of the membrane after staining it with fluorescent dyes³⁰⁻³². Small changes, of the order of one part in 10^4 to 10^6 , can be recorded in each case, but disappointingly they are either strictly voltage-dependent effects, or current-dependent effects which only arise when diffusion limitations cause a temporary accumulation of ions in the space between the Schwann cell and the membrane. No changes satisfactorily proved to be directly linked with membrane conductance, which might have thrown light on the molecular nature of the events in the ionic channels, have yet been reported.

There are three components of the change in optical retardation of the membrane in squid giant axons, which have relaxation times ranging from a few μs to about 20 ms²⁷. One of them, probably the fastest, can be attributed to an alteration in membrane thickness under the influence of the compressional force exerted by the attraction between the opposite charges on either side of the membrane. The other two are thought to arise from Kerr-type effects in membrane constituents that have still to be identified or even located, except that the slowest change (the "rebound" in our²⁷ terminology) probably takes place at the outside of the membrane, because its size is dependent on the external calcium concentration. All three components of the retardation response must be regarded as involving some kind of conformational change in the membrane, because of their relatively large relaxation times. As Cone³³ has recently shown, a molecule as large as rhodopsin can rotate in the membrane with a time constant of no more than 20 μs if it is free to do so. It is unlikely that we have to deal with the reorientation in the electric field of more than part of a molecule; but at the same time we have to account for a time constant up to three orders of magnitude greater. The best analogy is therefore with the conformational change in an allosteric enzyme, the relaxation time of which may well be in the millisecond range³⁴.

The several voltage-dependent light scattering changes²⁹ may either be another manifestation of the electrostriction effect, or may again involve conformational changes in unidentified membrane components. Although a detailed study of their relaxation times might enable some more positive identifications to be made, light scattering studies do not now seem to offer an especially promising approach.

The use of fluorescent probes offers greater possibilities, but as far as nerve is concerned, the results reported so far seem to relate to simple movements of the charged dye molecules under the influence of the electric field at the membrane surface, rather than to ionic conductance. Such probes may therefore turn out to be more useful as indicators of membrane potential and other changes in organelles like mitochondria³⁵ where microelectrodes cannot be inserted to make direct measurements, than in studies of the special properties of excitable membranes. Also the relatively large bulk of most fluorescent groups will always tend to raise a doubt as to whether the probing molecule is not unduly perturbing the structure into which it is introduced. However, Davila *et al.*³² have reported interesting variations in the relaxation time and voltage dependence of the fluorescence response for different dyes recorded with squid giant axons, analysis of which should ultimately yield useful information about the detailed structure of the surface of the membrane. Another hopeful line of attack is to use polarized ultraviolet illumination³⁶. Finally, if fluorescent groups could be attached to blocking agents like tetrodotoxin or quaternary ammonium ions, something further could be learnt about the nature of the sodium and potassium

channels, though to incorporate the fluorescent group without at the same time destroying the high affinity of the agent for its target may not be at all easy.

This new evidence on the occurrence of potential-dependent structural changes in the membrane does not conflict with Cole and Curtis's³ failure to detect any change in the membrane capacity, because the expected²⁷ alteration in membrane thickness is only 0.03 to 0.5%, well below their lower limit of 2%. Nor does it argue against the supposition that a rise in ionic conductance involves a conformational change at each active site, for the numerical disparity between the number of sites and the number of phospholipid and other molecules that make up the membrane is so great. The scales are in fact heavily weighted against obtaining direct information about the excitability mechanism from any of the physical techniques considered so far.

Gating Mechanism

An important characteristic of the sodium and potassium channels is the steepness of the quantitative relationship between conductance and voltage. In any physical system in which current is controlled by the displacement of singly charged particles or dipoles, application of the Boltzmann principle leads to the conclusion that an e -fold change in conductance must be associated with a potential change equal to kT/e , which at room temperature is about 25 mV. In the squid giant axon, however, an e -fold increase in sodium conductance can be brought about by a change in membrane potential of less than 5 mV, and the figure for potassium conductance is not very different³⁷. This implies that the alterations in permeability must depend on the simultaneous movement of a number of charges. These might all be located on one mobile particle, or there might be a cooperative displacement of several particles each with a smaller charge. If all the gating particles were initially located on one side of the membrane, then the opening of the gate on reduction of the membrane potential would be perceived as a pulse of current equal to the product of the number of channels and the amount of charge transferred at each one. Chandler and Meves³⁸ made some measurements of this so-called "gating current" in squid axons. In these experiments the ionic current was reduced to a minimum by substituting choline for sodium outside and perfusing the axon with a low ionic strength solution containing only a small quantity of rubidium chloride. The capacity current for a hyperpolarizing voltage-clamp pulse of equal size to the test depolarizing pulse was then subtracted. This left the gating current, which turned out to be below the limit that could reliably have been detected—estimated at 10^{-8} coulombs cm^{-2} . There are two possible explanations for this negative finding. One is just that greater instrumental sensitivity is needed. If the number of sodium channels is the same in squid axons as it is in lobster nerve¹⁷, the gating current would be 36×10^8 sites cm^{-2} multiplied by $5 \div (6.3 \times 10^{18})$ coulombs/site = 0.3×10^{-8} coulombs cm^{-2} . Alternatively, the passage of gating current may not be restricted to depolarizing pulses as tacitly assumed above. It would not be hard to imagine a different type of system in which the amount of work done on the membrane by the electric field, which may be equated with the gating current, would be the same for pulses in either direction, even though the gate was only opened by a depolarization. In this case the gating current would tend to be inextricably merged with the 5×10^{-8} coulombs cm^{-2} of capacity current corresponding to a 50 mV voltage-clamp pulse. Thus more accurate determinations of the gating current may do no more than establish the symmetry of the system.

Although the concept of a multi-charged mobile sodium carrier cannot be finally dismissed, it seems more likely that the gating mechanism is based on a cooperative phenomenon in which quite a large number of membrane mole-

cules may play a part. The precise form of the cooperation is still unclear, but one simple possibility is for the gate to be operated mechanically via the pressure exerted by the electrical field. But simplicity is relative; and not only does the opening of the gate by a depolarization have to be explained, but also all the complications of the subsequent inactivation and reactivation of the mechanism. If a phase transition is to be invoked, it must be a complex one with three different states, in which several types of molecule participate.

Selectivity of the Ionic Channels

It seems clear^{23,39} that the selectivity of the channels must depend fundamentally on the manner in which the cations interact with fixed anionic or hydrogen-bonding sites so as to traverse the membrane without some at least of the water molecules from their inner coordination sphere. Although the free energies both of solvation and of interaction with fixed sites will increase monotonically as the ionic radius decreases, the difference will not in general be monotonic. Subtleties of the architecture of the channels in different tissues can therefore be expected to lead to the well-known series of selectivity sequences³⁹. In the squid giant axon, for example, the sequence is $\text{Li}=\text{Na}>\text{K}>\text{Rb}>\text{Cs}$ for the sodium channel³⁸, and $\text{K}>\text{Rb}>\text{Na},\text{Cs}$ for the potassium channel⁴⁰. From measurements of the currents flowing through the sodium channels at the nodes of Ranvier in frog nerves, Hille⁴¹ has concluded that although some small organic cations like hydrazine and hydroxylamine can cross the membrane almost as easily as sodium or lithium, others that are no larger such as methylamine cannot. He suggests that while size is one of the factors involved, the channel behaving as if it were about 3 by 5 Å in cross section, permeant cations must also be capable of making hydrogen bonds to strategically placed oxygen atoms in the channel wall. Hence there is a selection against methyl groups. The same discrimination against methylamine is exhibited by the squid giant axon. In some other species either the structure of the sodium channels is different or they are largely replaced by another type of calcium-selective channel, for whereas in squid axons the permeability coefficient for calcium ions is of the order of one hundredth of that for sodium ions⁴², the early conductance change in barnacle muscle fibres allows calcium to enter in preference to sodium⁴³. The potassium channels seem to be constructed in a similar way, but are definitely narrower⁴⁴.

From his elegant studies of the interaction of quaternary ammonium derivatives with the potassium channels in squid axons, Armstrong⁴⁵ has proposed that they are gramicidin-like tunnels with a large mouth at their inner end. This mouth can accommodate a hydrated potassium ion or a tetraethylammonium ion, but in agreement with Hille⁴⁴, only a dehydrated potassium ion can pass through the tunnel itself. It seems probable that the entrance to the sodium channels from the outside tapers in a similar fashion, and that substances like tetrodotoxin owe their high affinity to an exact fit into it. A detailed comparison of the shapes of tetrodotoxin and saxitoxin molecules, each of which has a positively charged guanidinium group at one end, but otherwise quite different formulae, might enable a precise specification to be drawn up for the structure of the sodium channel mouth.

These rather exacting structural requirements provide support for the other evidence that the sodium and potassium channels are permanently in being, rather than—as has sometimes been suggested—having a strictly transient existence during some kind of phase transition. It remains unknown whether the gating mechanism operates in series with and to some extent independently of the selectivity mechanism, or whether it is a conformational change in the selective channel itself that opens and shuts it.

Action of Calcium

Another important influence on the excitation mechanism is that of external calcium. Voltage-clamp studies⁴⁶ have shown that in peripheral nerve the main effect of changes in $[Ca^{2+}]_o$ is to shift the curves relating sodium and potassium conductance to membrane potential sideways along the voltage axis without altering their shape or the magnitude of the peak conductances. An attractive way of explaining this effect, but one which had to be rejected at an early stage, was that calcium ions might simply block the sodium channels by occupying critical sites on the membrane from which they would be released by a reduction in potential. Unfortunately a quantitative consideration of the shift in potential corresponding to a given change in calcium concentration showed that this hypothesis was untenable. Instead, the action of calcium and the other alkaline earth elements seems to involve their adsorption at the outer edge of the membrane in such a manner that the electric field controlling the sodium and potassium channels is modified, although the over-all potential difference between inside and outside is unaltered. The fact, among others, that squid axons can be exposed to solutions containing magnesium but no calcium for many hours without any marked change in their sodium conductance (unpublished observations of E. Rojas and R. D. K.) makes it difficult to attribute any more specific role to calcium.

What has just been said is not, of course, true for several other membrane mechanisms. There is ample evidence that calcium has a special role in the activation of muscular contraction, in the release of acetylcholine and other chemical transmitters at nerve endings, and probably in other processes as well, but these fall outside the scope of this article.

Isolation of Receptor Sites

It has been reported by de Robertis⁴⁷ that a proteolipid thought to be the acetylcholine receptor can be extracted from brain tissue and electric organ in organic solvent mixtures and yet retain its specific binding properties. From equilibrium dialysis studies of a particulate preparation from *Torpedo* electric organ, Eldefrawi *et al.*⁴⁸ have questioned the validity of this identification; and there are other reasons (unpublished observations of S. R. Levinson) for doubting whether the chloroform/methanol-soluble protein is still genuinely labelled. But less drastic extraction procedures can certainly be used without detaching specific blocking agents from acetylcholine receptor sites. Thus Miledi and Potter⁴⁹ have succeeded in solubilizing receptors attached to ¹²⁵I-labelled α -bungarotoxin by treating membranes from electric organ and frog muscle with a non-ionic detergent, which presumably helps to form a protective collar around the hydrophobic portions of the molecule, so that it remains structurally intact in an aqueous environment. This kind of technique offers the most promising approach for learning more about the chemical composition of the ionic channels in electrically excitable membranes. Again on a basis of equilibrium dialysis studies, Denburg *et al.*⁵⁰ have reported the presence in lobster nerves of cholinergic receptors that bind nicotine and procaine. But their suggestion that these may be a component common to both the sodium and potassium channels does not seem to be well founded, partly because they are too numerous by a factor of twenty, and partly because there is no good evidence that acetylcholine is directly involved in nervous conduction. It would, all the same, be interesting to establish what their role in these nerves really is.

- ⁴ Fettiplace, R., Andrews, D. M., and Haydon, D. A., *J. Membrane Biol.*, **5**, 277 (1971).
- ⁵ Singer, S. J., *Structure and Function of Biological Membranes*, 145 (Academic Press, New York, 1971).
- ⁶ Singer, S. J., and Nicolson, G. L., *Science*, **175**, 720 (1972).
- ⁷ Frye, C. D., and Edidin, M., *J. Cell Sci.*, **7**, 313 (1970).
- ⁸ Gitler, C., *Ann. Rev. Biophys. Bioengin.*, **1** (1972).
- ⁹ McConnell, H. M., and McFarland, B. G., *Quart. Rev. Biophys.*, **3**, 91 (1970).
- ¹⁰ Radda, G. K., *Current Topics in Bioenergetics*, **4**, 81 (Academic Press, New York, 1971).
- ¹¹ Finer, E. G., Flook, A. G., and Hansen, H., *Biochim. Biophys. Acta*, **260**, 49 (1972).
- ¹² Krasne, S., Eisenman, G., and Szabo, G., *Science*, **174**, 412 (1971).
- ¹³ Howarth, J. V., Keynes, R. D., and Ritchie, J. M., *J. Physiol.*, **194**, 745 (1968).
- ¹⁴ Keynes, R. D., *The Neurosciences: Second Study Program*, 707 (Rockefeller University Press, New York, 1970).
- ¹⁵ Hille, B., *Progr. Biophys. Mol. Biol.*, **21**, 1 (1970).
- ¹⁶ Moore, J. W., Narahashi, T., and Shaw, T. I., *J. Physiol.*, **188**, 99 (1967).
- ¹⁷ Keynes, R. D., Ritchie, J. M., and Rojas, E., *J. Physiol.*, **213**, 235 (1971).
- ¹⁸ Colquhoun, D., Henderson, R., and Ritchie, J. M., *J. Physiol.*, **224**, 25P (1972).
- ¹⁹ Wagner, H. H., Schwarz, J. R., and Ulbricht, W., *Pflügers Arch. Ges. Physiol.*, **332**, R 68 (1972).
- ²⁰ Cuervo, L. A., and Adelman, W. J., *J. Gen. Physiol.*, **55**, 309 (1970).
- ²¹ Katz, B., and Miledi, R., *Nature New Biology*, **232**, 124 (1971).
- ²² Gordon, L. G. M., and Haydon, D. A., *Biochim. Biophys. Acta*, **255**, 1014 (1972).
- ²³ Eigen, M., and Winkler, R., *The Neurosciences: Second Study Program*, 685 (Rockefeller University Press, New York, 1970).
- ²⁴ Cohen, L. B., Keynes, R. D., and Hille, B., *Nature*, **218**, 438 (1968).
- ²⁵ Cohen, L. B., Hille, B., and Keynes, R. D., *J. Physiol.*, **203**, 489 (1969).
- ²⁶ Cohen, L. B., Hille, B., and Keynes, R. D., *J. Physiol.*, **211**, 495 (1970).
- ²⁷ Cohen, L. B., Hille, B., Keynes, R. D., Landowne, D., and Rojas, E., *J. Physiol.*, **218**, 205 (1971).
- ²⁸ Cohen, L. B., and Keynes, R. D., *J. Physiol.*, **212**, 259 (1971).
- ²⁹ Cohen, L. B., Keynes, R. D., and Landowne, D., *J. Physiol.*, **224**, 701 (1972).
- ³⁰ Tasaki, I., Watanabe, A., Sandlin, R., and Carnay, L., *Proc. US Nat. Acad. Sci.*, **61**, 883 (1968).
- ³¹ Conti, F., and Tasaki, I., *Science*, **169**, 1322 (1970).
- ³² Davila, H. V., Cohen, L. B., and Waggoner, A. S., *Biophys. Soc. Abstr. SaAM-E13* (1972).
- ³³ Cone, R. A., *Nature New Biology*, **236**, 39 (1972).
- ³⁴ Hammes, G. G., *Adv. Protein Chem.*, **23**, 1 (1968).
- ³⁵ Chance, B., *Proc. US Nat. Acad. Sci.*, **67**, 560 (1970).
- ³⁶ Hallett, M., Tasaki, I., Schneider, A., and Carbone, E., *Biophys. Soc. Abstr.*, SaAM-F7 (1972).
- ³⁷ Hodgkin, A. L., *The Conduction of the Nervous Impulse* (Liverpool University Press, 1964).
- ³⁸ Chandler, W. K., and Meves, H., *J. Physiol.*, **180**, 788 (1965).
- ³⁹ Eisenman, G., *Biophys. J. (Suppl.)*, **2**, 259 (1962).
- ⁴⁰ Pickard, W. F., Lettvin, J. Y., Moore, J. W., Takata, M., Pooler, J., and Bernstein, T., *Proc. US Nat. Acad. Sci.*, **52**, 1177 (1964).
- ⁴¹ Hille, B., *J. Gen. Physiol.*, **58**, 599 (1971).
- ⁴² Baker, P. F., Hodgkin, A. L., and Ridgway, E. B., *J. Physiol.*, **218**, 709 (1971).
- ⁴³ Hagiwara, S., Hayashi, H., and Takahashi, K., *J. Physiol.*, **205**, 115 (1969).
- ⁴⁴ Hille, B., *Biophys. Soc. Abstr.*, SaAM-E12 (1972).
- ⁴⁵ Armstrong, C. M., *J. Gen. Physiol.*, **58**, 413 (1971).
- ⁴⁶ Frankenhaeuser, B., and Hodgkin, A. L., *J. Physiol.*, **137**, 218 (1957).
- ⁴⁷ Robertis, E. de, *Science*, **171**, 963 (1971).
- ⁴⁸ Eldefrawi, M. E., Britten, A. G., and Eldefrawi, A. T., *Science*, **173**, 338 (1971).
- ⁴⁹ Miledi, R., and Potter, L. T., *Nature*, **233**, 599 (1971).
- ⁵⁰ Denburg, J. L., Eldefrawi, M. E., and O'Brien, R. D., *Proc. US Nat. Acad. Sci.*, **69**, 177 (1972).

¹ Hodgkin, A. L., and Huxley, A. F., *J. Physiol.*, **117**, 500 (1952).

² Hodgkin, A. L., and Horowicz, P., *J. Physiol.*, **148**, 127 (1959).

³ Cole, K. S., and Curtis, H. J., *J. Gen. Physiol.*, **22**, 649 (1939).

Prostaglandins and Aspirin

V. R. PICKLES

Physiology Department, University College, Cardiff

This article has the same title as one published in *Nature* about fourteen months ago¹. Its purpose is to re-examine a thesis put forward then, and to mention some of the recent developments.

THIS article is not intended primarily for pharmacologists and physiologists, so it may be helpful to sketch in the background first. Prostaglandins are certain fatty-acid derivatives that are biologically highly active in various ways and are produced by many organs. They are still described (as in a recent undergraduate textbook) in such terms as "newly discovered substances of unknown function", but this is somewhat misleading: the word "prostaglandin" was introduced in 1935, at the same time as "progesterone"^{2,3}; the chemical nature of the primary prostaglandins has been known for a decade, and several particular physiological functions are well established. I shall return later to the question of their general function. The previous article¹ lists several reviews and published symposia on prostaglandins, and in the past few months two more books^{4,5} and three further reviews⁶⁻⁸ have appeared.

Anti-inflammatory Drugs

Aspirin is one of the non-steroidal anti-inflammatory drugs. Both Collier's article¹ and this one are concerned with the group as a whole and perhaps particularly with indomethacin, which is used in the treatment of such conditions as rheumatoid arthritis. The earlier article first traced the background to the discovery by Vane and his colleagues that indomethacin, aspirin and sodium salicylate can inhibit the biosynthesis of prostaglandins, which, moreover, they do *in vitro* in concentrations likely to be reached in the body by therapeutic doses. This discovery immediately answered some of the outstanding problems in the mechanism of action of these anti-inflammatory drugs: for example, the antipyretic action of aspirin can be understood as following the inhibition of synthesis, in the temperature regulating region of the brain, of prostaglandin E₁ which is a potent pyrogen⁹.

Recent reports have confirmed the original discovery. For example, a dose of only eight aspirin tablets a day temporarily reduced the amount of various prostaglandins in the seminal fluid of young male volunteers¹⁰; indomethacin or aspirin abolished the release of prostaglandins from the spleen when it was stimulated to contract¹¹. Not only is the synthesis of prostaglandins themselves affected, but also that of "rabbit-aorta-contracting substance" (RCS)¹² which is likely to be a normal intermediate in the formation of prostaglandins. In a review of his group's work on this topic, Vane has proposed the concept of a "cascade" of physiologically active substances, of which the prostaglandins are the end products¹³.

In 1969, Collier had presented a detailed pharmacological analysis¹⁴ of the actions of aspirin, developing the concept that drugs of this nature inhibit the "defensive reactions" of the body. His more recent article¹ examined the evidence then available that it is one function ("the" function?) of prostaglandins to mediate some of these defensive reactions, and that the anti-inflammatory drugs act in part either by depressing the biosynthesis of the prostaglandins, or by preventing the action of prostaglandins already formed. Certain difficulties were

recognized in this thesis, and in some instances it was possible to propose ways of escape from them. One of the principal difficulties may, however, lie in the general idea of "defensive reactions".

As we normally communicate by means of words rather than the symbols of mathematics, physics and chemistry, it is difficult to avoid introducing into biology concepts that are derived from ordinary social experience. To take a trivial example, innumerable examination questions ask candidates to describe the role of (or the part played by) such-and-such in the body; the candidate knows what response is expected, without pausing to think that the metaphor is that "all the body's a stage, the autacoids and hormones merely players". Although it is often true that these too have their exits and their entrances (origin and fates) and may in their time play many parts, one knows not to push the analogy too far. Now I submit that the term "defensive reaction" is likewise a phrase derived from social interaction, and if its application to particular facts of physiology or pathology causes difficulty, then the metaphor is to that extent inappropriate, not the facts necessarily wrong. Two examples from the earlier article¹ will illustrate the problem. The bronchoconstriction in asthma may defend the delicate alveolar surface, yet no catastrophic invasion of the body follows its relief by pharmacological means. Again, implantation of the ovum has been likened to inflammation (a "defensive" function), which led to the suggestion that this might "explain" the presence of prostaglandin in semen. If seminal plasma drawn into the cavity of the uterus has any effect there, this effect would seem unlikely to persist for some days until the fertilized ovum reached the cavity of the uterus and began to be implanted. At least two possible (and not incompatible) functions of seminal-fluid prostaglandin have been proposed and not disproved: these were von Euler's¹⁵ early ideas that they might facilitate the contractions of smooth muscle in the male genital tract during orgasm, and Eliasson's¹⁶ that their effects on the female genital tract would help sperm transport. A further physiological effect of prostaglandins on the uterus is given by the prostaglandins produced by the endometrium itself, which stimulate the expulsive contractions of the myometrium during menstruation¹⁷—this "menstrual stimulant"¹⁸⁻²⁰ action being the physiological basis of the early abortifacient action of prostaglandins that is now arousing so much interest. Both menstruation and abortion (natural or induced) may, if one finds it helpful, be thought of as defensive reactions mediated by prostaglandins. To the physiologist, however, it may seem more useful to think of the "menstrual stimulant" function of endometrial prostaglandins as an example of a positive feedback loop, which could ultimately be described in precise cybernetic terms.

Even more generally, one may ask how likely it is that a group of substances defined in chemical terms should neatly fit a broad functional (teleological?) classification. At least, a lack of perfect fit seems no cause for worry.

Other Uses

Whatever one's opinions on such matters may be, it is of interest to ask whether further clinical applications of some anti-inflammatory drugs may be suggested by their anti-prostaglandin actions. Such suggestions include the reduction of fluid loss in cholera²¹, the prevention of unwanted abortion, and the treatment of raised intra-ocular pressure in some diseases¹³ (this deriving from the pioneer work of Ambache on irin²²); it has also been suggested that anti-inflammatory drugs might have untoward effects in preventing intra-uterine con-

traceptive devices from exerting their effect, or in retarding the healing of wounds¹³. Work is known to be in progress on some of these possibilities.

One of the anti-inflammatory drugs could possibly prove of value in dysmenorrhoea, which may be associated with abnormal prostaglandin action²³, especially as aspirin has commonly been recommended for its simple treatment. (Alcohol, also known to inhibit myometrial responses to prostaglandins, has also been stated to be effective, but for obvious reasons unsuitable for routine treatment!).

The fact that effects such as these have not already been clearly described may only mean that they were not looked for. Even the simple observation that aspirin by mouth diminishes the prostaglandin content of human semen could have been made, in a rough way, at any time since the early 1930s, but in fact was not made until interest was aroused by the fundamental observations *in vitro*.

It is tempting to speculate that notable further discoveries will be made about the functions of prostaglandins in the central nervous system. Indomethacin, in blood concentrations of the same order as those effective against prostaglandin synthesis *in vitro*, causes disturbances of equilibrium, feelings of unreality or dissociation, and intellectual impairment, as well as headache, which presumably may be of meningeal origin²⁴. It is not known whether any of these symptoms can be attributed to an effect on prostaglandin synthesis.

¹ Collier, H. O. J., *Nature*, **232**, 17 (1971).

- ² Euler, U. S. von, *Klin. Wschr.*, **14**, 1182 (1935).
³ Allen, W. M., Butenandt, A., Corner, G. W., and Slotta, K. H., *ibid.* (1935).
⁴ Horton, E. W., *Prostaglandins* (William Heinemann Medical Books, London, and Springer-Verlag, Berlin, Heidelberg and New York, 1972).
⁵ *The Prostaglandins: Progress in Research* (edit. by Karim, S. M. M.) (Medical and Technical Publishing Co., Oxford and Lancaster, 1972).
⁶ Pike, J. E., *Scientific American*, **84** (November 1971).
⁷ Hinman, J. W., *Ann. Rev. Biochem.*, **41** (in the press).
⁸ Weeks, J. R., *Ann. Rev. Pharmacol.*, **12**, 317 (1972).
⁹ Milton, A. S., and Wendtland, S., *J. Physiol.*, **207**, 76P (1970).
¹⁰ Collier, J. G., and Flower, R. J., *Lancet*, **ii**, 852 (1971).
¹¹ Ferreira, S. H., Moncada, S., and Vane, J. R., *Nature New Biology*, **231**, 237 (1971).
¹² Gryglewski, R., and Vane, J. R., *Brit. J. Pharmacol.*, **43**, 420P (1971).
¹³ Vane, J. R., *Hospital Practice*, 61 (March 1972).
¹⁴ Collier, H. O. J., *Adv. Pharmacol. Chemother.*, **7**, 333 (1969).
¹⁵ Euler, U. S. von, *J. Physiol., Lond.*, **88**, 213 (1936).
¹⁶ Eliasson, R., *Acta Physiol. Scand.*, **46**, Suppl., 158 (1959).
¹⁷ Eglinton, G., Raphael, R. A., Smith, G. N., Hall, W. J., and Pickles, V. R., *Nature*, **200**, 960 (1963).
¹⁸ Pickles, V. R., *Nature*, **180**, 1198 (1957).
¹⁹ Chambers, P. L., and Pickles, V. R., *J. Physiol., Lond.*, **144**, 68 (1958).
²⁰ Clitheroe, H. J., and Pickles, V. R., *J. Physiol., Lond.*, **156**, 255 (1961).
²¹ Bennett, A., *Nature*, **231**, 536 (1971).
²² Ambache, N., *J. Physiol., Lond.*, **135**, 114 (1957).
²³ Pickles, V. R., Hall, W. J., Best, F. A., and Smith, G. N., *J. Obstet. Gynaec. Brit. Commonw.*, **72**, 185 (1965).
²⁴ Rothermich, N. O., *J. Amer. Med. Ass.*, **195**, 531 (1966).

Developments in the Physics of Elementary Particles

A. ASHMORE

Daresbury Nuclear Physics Laboratory, Keckwick Lane, Daresbury, Warrington WA4 4AD

There are signs of a breakthrough in the understanding of the structure and interactions of elementary particles, but no one can predict precisely how it will come. The best guarantee of early success is to continue to attack the subject on all fronts.

EXPERIMENTS on the intersecting storage rings (ISR) at CERN have now been running for a little more than a year. For the proton-proton interaction these give access to energies in the centre of mass system higher than can be obtained from any accelerator, existing or under construction. The range of momentum for each of the colliding beams is 10 to 30 GeV/c; at the lower end this overlaps with what will be available at the National Accelerator Laboratory in the United States and at the new super proton synchrotron (SPS) at CERN, and at the upper end it corresponds to an energy above 1,000 GeV for a proton beam impinging on a stationary target. The rate of collisions in the intersecting regions is, however, several orders of magnitude lower than can be achieved with a proton beam extracted from an accelerator and a stationary target. So far the experiments have been such that this is not a serious limitation, for they have been confined to small transfers of momentum where the cross sections are relatively high. The possibility of finding quarks may, however, be precluded by the restriction on the rate of interactions rather than the available energy. No quarks have yet been observed but the searches for them are still in their early stages.

Elastic Scattering

The forward peak in elastic scattering has been studied at the

ISR in two experiments with statistical accuracy at least as good as experiments at lower energies. One of the difficulties with colliding beams is to derive absolute cross sections because the rate of interactions varies inversely as the effective height of the intersecting region. This can be determined from the change in rate as the two beams are displaced relative to one another in a vertical direction but the accuracy of this method is limited to about 5%. It was already known from experiments with a stationary target at CERN and Serpukhov that this forward peak shows little change for momenta of the proton beam greater than about 10 GeV/c. This corresponds to a value for the squared centre of mass energy s of 21 GeV². Experiments at Serpukhov have been made up to a momentum of 70 GeV/c corresponding to $s = 134$ GeV². The peak can be fitted with a straight line in a plot of the cross section on a logarithmic scale against the squared momentum transfer t , providing the fitting is limited to values of t below about 0.3 (GeV/c)². Although this value is very low compared with the momentum of the incident proton, the forward peak is so sharp that most of the cross section for elastic scattering is within this limit. The slope of the straight line is found to increase slowly with energy, from 9.0 to 11.5 (GeV/c)⁻² between the values of s mentioned above. This is the effect known as shrinkage, indicating that the proton is not behaving simply as an absorbing sphere of fixed size.

The data from the ISR cover a range of values of s from about 500 to 3,000 GeV², and the striking feature is a sharp break to a line of lower slope above $t = 0.1$ (GeV/c)². If it is assumed that the initial steeper line is to be compared with the data at lower energies then the shrinkage is still continuing up to the highest energy but more slowly, suggesting that there may be little further shrinkage. The slope reaches a value of 12.5 (GeV/c)⁻² at the highest energy. The second line shows no evidence of shrinkage and has a slope, averaged over the energies of observation, of 10.7 (GeV/c)⁻². It may be that some of the experiments at lower energies should now be repeated with

increased precision to trace the origin of the break in slope, which certainly raises new questions about the structure of the proton and its behaviour in the limit of very high energies.

In the high energy limit it was suggested by Pomeranchuk in 1956 that the total cross section for the interaction between two particles will become constant, independent of energy. With a secondary or extracted beam the total cross section for the interaction of a particle with a proton can be determined very accurately from the loss of particles out of the beam when it traverses a target of liquid hydrogen. Results from transmission measurements of this kind are consistent with a constant total cross section for proton-proton scattering at a value just below 39 mb ($1 \text{ mb} = 10^{-31} \text{ m}^2$) over the range of beam momentum 20 to 60 GeV/c. Such measurements are not possible at the ISR and values of the total cross section have so far only been deduced from the elastic scattering cross section in the forward direction using the optical theorem. This is known from measurements at Serpukhov to be correct within 1% and there is reason to suppose the difference would be less at higher energies. As mentioned above, the absolute accuracy of cross sections at the ISR is limited to about 5%. Within this error the total cross section is independent of energy with a value of $37 \pm 1 \text{ mb}$, which is not significantly different from the value from transmission measurements. Work is in progress to improve the accuracy of measurement of the effective height of the beam and to make a direct measurement of the total cross section by having detectors covering as far as possible the complete angular range.

A further indication of the extent to which the high energy limit has been reached is given by measurements of the total cross sections of particles and their antiparticles (π^+ and π^- , K^+ and K^- , p and $\bar{p}$) at Serpukhov. The cross sections are expected to become equal and in fact the differences are decreasing according to a simple power law. Equality within the error will not be reached until momenta of 10,000 to 100,000 GeV/c are reached. This is clearly beyond the range of any conceivable accelerator but it will be interesting to see whether the present trends are continued to several hundred GeV. At 500 GeV/c, for example, the difference is expected to fall to about 0.5 mb for pions and kaons. An interesting feature of the results from Serpukhov is the rise of the total cross section for K^+p scattering above 20 GeV/c after being approximately constant from 2.0 to 20 GeV/c.

Inelastic Scattering

At high energies elastic scattering accounts for only about 20% of the total cross section, the rest being particle production with a great variety of final states. Cross sections for individual processes usually decrease rapidly with increasing energy but the variety of processes increases to maintain the overall cross section. If the production of a particular type of particle is observed, irrespective of what other particles are present in the final state, a simple scaling behaviour is found which is independent of the energy at which the experiment is made. Such behaviour was discovered in the deep inelastic scattering of high energy electrons and has opened up a most promising avenue towards a deeper understanding of the structure and interactions of elementary particles. Data for such inclusive reactions have been obtained at the ISR for pions, kaons, protons and anti-protons. Scaling seems to hold over the range of energies at the ISR, $s = 500$ to $3,000 \text{ GeV}^2$. In the case of pions, protons and probably positive kaons this extends down to the region of experiments with stationary targets at a beam momentum around 20 GeV/c, corresponding to $s = 39 \text{ GeV}^2$. For negative kaons and antiprotons the cross sections are larger at the ISR than at the lower energy by respective factors of about 2 and 7. The results confirm the exponential fall in yield with increasing transverse momentum of the produced particle observed at lower energies. The rate of fall depends on the type of particle but is in the region of an order of magnitude for an increase in transverse momentum of 0.5 GeV/c. It will

readily be appreciated that there are difficulties in making observations for large values of via transverse momentum.

Probing with Electrons

The use of the electromagnetic current to probe the structure of elementary particles is well established in the case of the proton and neutron through the elastic scattering of high energy electrons; the interaction can be visualized as taking place via the exchange of a virtual photon. This technique has been extended in recent years to inelastic scattering in which strongly interacting particles are produced. Experiments in which only the scattered electron is observed show a continuum in momentum which persists at large values of momentum transfer by contrast with the elastic scattering. This is the deep inelastic scattering mentioned above. The dependence on momentum transfer is suggestive of scattering from a point rather than from an extensive structure, and thus a substructure within the nucleon. The word "parton" has been coined to describe these hypothetical constituents of the nucleon and speculation has included the possibility that they may be the elusive quarks. There seems, however, to be some restriction that prevents them from materializing, and it is only the familiar particles that emerge after the interaction. The technique of inelastic electron scattering is also being used to explore the structure of particles other than the nucleon by observing coincidences between the scattered electron and one of the particles produced. In this way, for example, the structures of the pion and the baryon resonances have been investigated. There is increasing evidence that the strongly interacting particles have a characteristic size of about 1 fm (10^{-15} m) and it is their composite structure on this scale which determines the features of individual reactions. The simple scaling behaviour of deep inelastic scattering and inclusive reactions is then correspondingly associated with incoherent scattering from the point-like constituents.

It is clearly important to extend the studies of deep inelastic scattering to higher energies, as can be done by using the secondary beams of electrons and muons from the new accelerators at NAL and CERN. An associated phenomenon is that the production of strongly interacting particles by colliding beams of positive and negative electrons has been found to persist at the highest available energy of 1.5 GeV for each of the colliding beams at Frascati. A similar facility has recently been brought into operation at the Stanford Linear Accelerator Centre (SLAC) with an initial capability of 2.4 GeV in energy and the expectation that higher currents will yield a greatly increased rate of interactions. Also the experimental programme has now started on the bypass at the Cambridge Electron Accelerator which has a similar capability in energy but a rate of interactions at present below that at Frascati.

Weak Studies Strong

It is equally interesting to explore strong interactions with the weak current, which is responsible for the decay of most of the "stable" particles—the long lived ones which do not decay through strong interactions. At high energies this requires beams of neutrinos, which are unique because they undergo only weak interactions. Such beams involve long flight paths for the decay of pions to produce the neutrinos and a large amount of shielding to remove muons. Also the cross sections are very low so that the data obtained are not comparable in quality with those from electron beams. An important step forward has, however, been taken with the commissioning of the new large heavy liquid bubble chamber, Gargamelle, at CERN. This gives an event about once a minute for interactions induced by neutrinos from the proton synchrotron, and there is the prospect of a five-fold increase when the 800 MeV booster, now being commissioned, is brought into operation. The use of neutrino beams will undoubtedly be an important part of the experimental programmes at NAL and the SPS at CERN.

Oceanic Fine Structure

H. CHARNOCK

National Institute of Oceanography, Wormley, Godalming, Surrey

Side scan sonar methods are increasing knowledge of the detailed topography of the bed of the ocean: a new long range device, which can produce sonographs at a range of 22 km, has located a tectonic plate boundary. Small scale fluctuations in the vertical profiles of temperature and salinity are exciting oceanographers, meteorologists and fluid dynamicists alike.

NEXT Christmas it will be a hundred years since HMS Challenger started her famous voyage of oceanographic exploration. She is often pictured "shortening sail to sound", a reminder that to measure the depth of the ocean using hemp or wire rope was a difficult and time consuming operation. The resulting sparsity of soundings gave the false impression that the seabed was relatively smooth; it was not until echo sounders were developed that the magnificent variability of the relief could be appreciated.

Even modern precision echo sounders have rather wide beam widths (typically 20° between half power points) with the result that the records of the small scale structure are distorted and can be misleading. A recent example¹ comes from the Mediterranean Ridge. Here conventional echo sounders produce records of characteristic undulations which resemble (and have been referred to as) cobblestones. Use of an echo sounder with a narrower beam shows that the seafloor is made up of offset steeply sloping blocks. They are probably a result of faulting.

Another technique which is being increasingly exploited—side scan sonar—involves transmitting sound in a fan-shaped beam, narrow fore and aft and broad in the vertical direction. The sound is reflected on rock outfalls and cliffs so that contrasts arise from differences in texture or slope. The echoes are received and displayed on board to form an image, or sono-

graph, as the ship steams on a steady course at a steady speed. Skilled interpretation of the resulting sonographs is needed: the effect can be compared with the heightened relief seen by headlights when driving over a rough road in the dark. The example of Fig. 1a (R. H. Belderson, N. H. Kenyon and J. B. Wilson, to be published) shows the scars made on the seabed by icebergs long since dissolved. Palaeocurrents seem to have been as variable as those of the present day.

To get side scan sonar records at longer ranges the equipment needed is much more complicated. To avoid refraction problems the sonar array has to be towed from the ship at a depth below that of the seasonal thermocline, typically at 75 m. Because the sound takes so long to travel to the maximum range and back (30 s for 22 km) imaging is only possible if the array is accurately towed in a straight line and in practice the array is electronically steered so that it can effectively listen to echoes from the last ping emitted.

The resulting vehicle, the Geological Long Range Inclined Asdic¹⁰, or GLORIA, is housed in a glass fibre towed body about 9 m long. It weighs about 7 tons in air and is difficult, sometimes dangerous, to launch from a ship. It is tied to its mother ship RRS Discovery, literally by a multicore cable for power, for signal generation and processing circuits, and for control, monitoring and so on. A control console in the ship processes the signal, using correlation techniques; the record is displayed on a conventional wet paper facsimile recorder.

In the past year the GLORIA system has been used to study the small scale relief of the Mediterranean Ridge (Stride, in preparation) for tracking shoals of herring in the Minches (Rusby, in preparation) and has provided unique information about the evolution of the North Atlantic².

During a survey from RRS Discovery the GLORIA record showed a continuous east to west reflector extending 400 km east of the Azores. This turned out to be a V-shaped valley, about 2 to 6 km wide and from 100 to 500 m deep. It seems clear that this is a fault (it is now called the GLORIA fault). It is thought to be the principal transcurrent fault in this area and to mark a boundary between the Eurasian and the African tectonic plates.

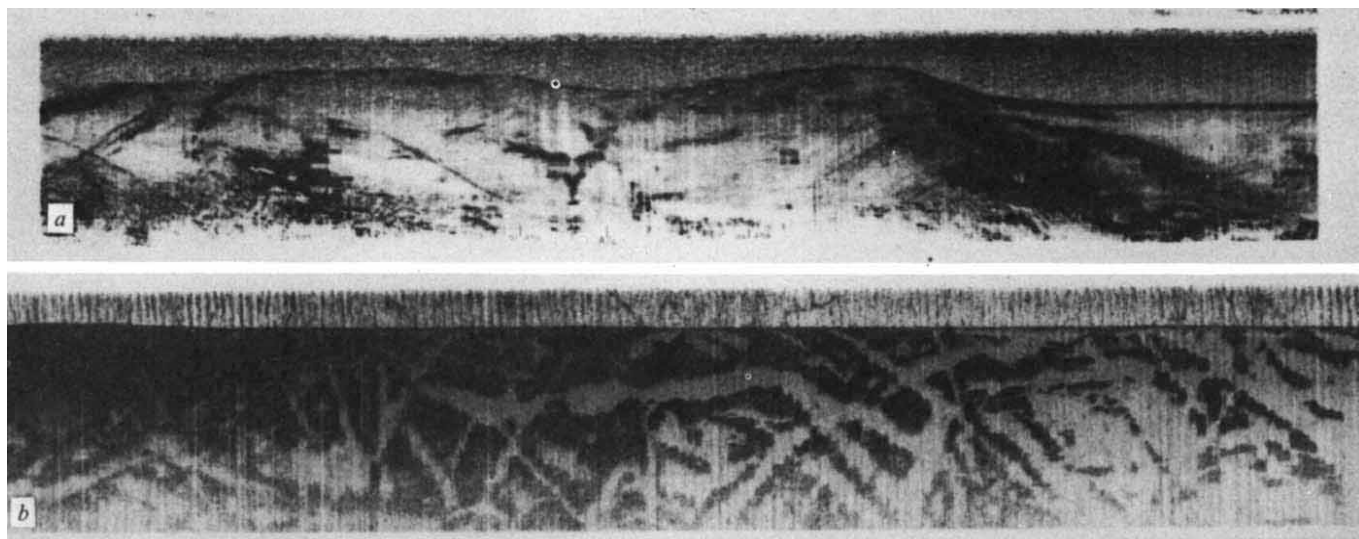


Fig. 1 a, Sonograph from 48 kHz transit sonar in depth of 150 m west of Shetland Islands. Range 550 m, distance along track 6.5 km. The dark tone is due to boulders and cobbles, the light tone to muddy sand which is infilling the furrows made by stranding icebergs (from ref. 2); **b**, GLORIA sonograph of part of the Palmer Ridge which divides Peake and Freen deeps in the North Atlantic. Palmer Ridge reaches about 3,000 m above the surrounding area, the depth of which varies from about 5,000 m to 2,500 m. Range 22 km, distance along track 70 km.

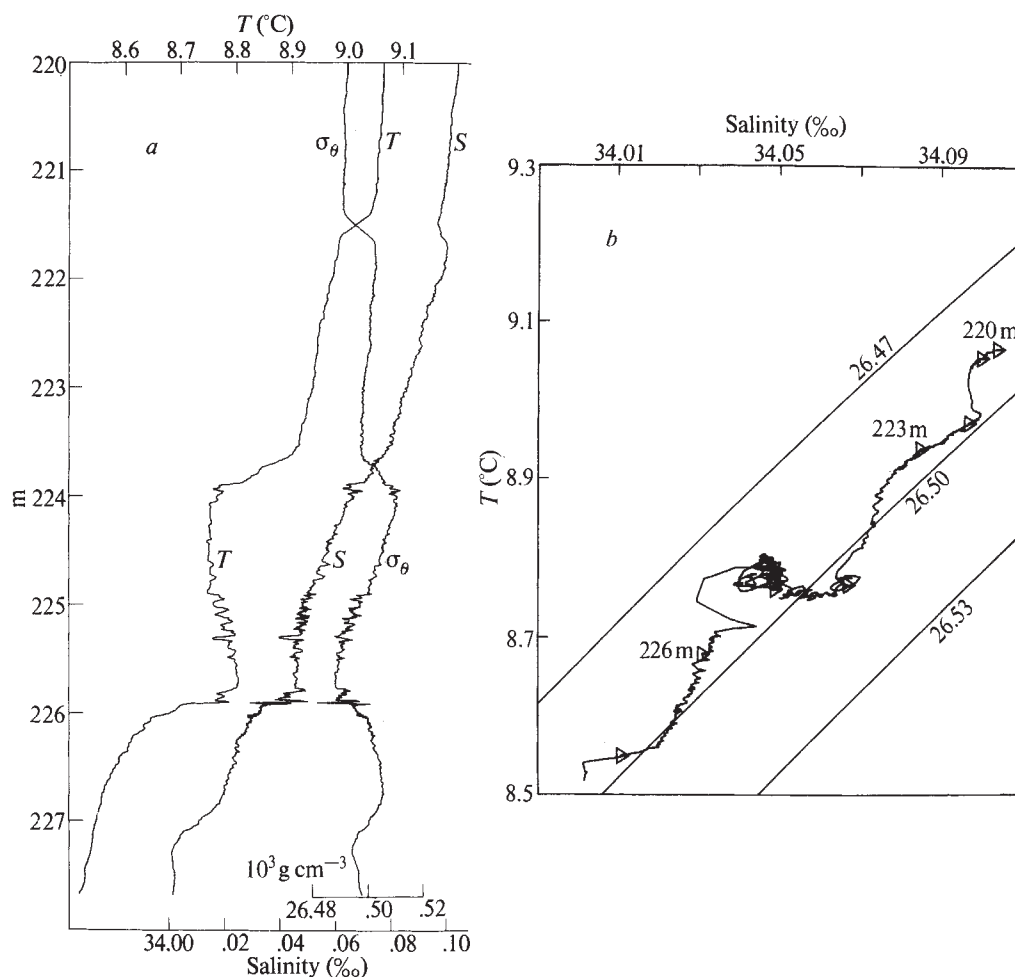


Fig. 2 *a*, Microstructure of temperature, salinity, and density between 224 m and 226 m over the San Diego Trough; *b*, temperature plotted against salinity from the data of Fig. 2*a*, with one point every 0.35 cm. (From ref. 4.)

Fine Structure of Temperature and Salinity

Physical oceanographers are devoting a lot of effort to studying fine structure in the water of the deep ocean. The classical method used mercury-in-glass thermometers on sampling bottles to estimate the temperature and salinity at a series of discrete depths (pressures). In this way the mean distribution of temperature, salinity and density in mid-ocean has been established. For temperature, for example, there is usually a well mixed layer near the surface, below which the temperature decreases approximately exponentially with a scale length of about 1 km.

The development of continuous *TSD* (temperature, salinity, depth) recorders has shown that the discrete sampling was obscuring much fine structure. The instruments consist of sensors for temperature (a resistance thermometer), salinity (a conductivity cell) and depth (a strain gauge pressure sensor) which modulate the frequencies of three oscillators. The combined signal is transmitted up a single core cable to the laboratory in the ship where it is decoded and displayed. The absolute accuracy obtained is hardly better than that of the classical method (0.01°C , 0.002°‰ , 5 debar) but the relative accuracy is higher. In really well mixed water the potential temperature is found to be constant within a few thousandths of a degree—a tribute to the instrument and to Knudsen (born 1871) who calculated the potential temperature correction many years ago.

The continuous records showed that the temperature and salinity did not vary smoothly with depth; they often had a series of steps that have been idealized as layers, where the temperature and salinity vary only slowly, and sheets where vertical gradients are much larger.

Commercially available *TSD* type instruments can resolve

structure on the scale of a few m—the limits being imposed by the size of the instrument, by the inertia of the sensors and by the mismatch in their time constants. Gregg and Cox³ have developed a high resolution probe which senses temperature and electrical conductivity (whence salinity) from a freely falling vehicle. It can sense features of less than 1 cm (smaller scale temperature variations are rapidly attenuated by thermal diffusion); features with vertical wavelength of more than 2 cm are fully resolved. Fig. 2 shows a sample of the record they have obtained.

Such observations of oceanic microstructure attract great interest, from fluid dynamicists and physical oceanographers alike. There are related structures in the atmosphere too, notably near the tropopause, so meteorologists are also concerned, especially with the clear air turbulence which is of growing concern as aircraft become bigger.

Much of the small scale structure in temperature and salinity has been found by Pingree⁴ to be highly positively correlated and density compensating. This argues that it could have been generated by local mixing, followed by horizontal spreading. Woods and Wiley⁵ observed the motion of dye tracers underwater and have shown that the layering is often associated with the (Kelvin-Helmholtz) instability of internal waves.

This raises the general problem of turbulence in free, stably stratified fluids: a valuable review of relevant laboratory experiments has been prepared by Thorpe⁶.

Another fascinating type of instability results from the difference in the rates of molecular diffusion for heat and salt in the ocean ($\kappa \approx 10^{-3}$, $D \approx 10^{-5} \text{ cm}^2 \text{ s}^{-1}$). A layer of warm salt water resting on a heavier layer of cold fresh is hydrostatically stable, but in the laboratory it produces long narrow convection cells, now known as salt fingers. These in turn produce a regular step-like vertical structure in temperature and salinity: struc-

tures resembling them have been observed in the ocean⁷. If the heavier lower layer is warm but salty the result is a series of sharp interfaces separating turbulent layers⁸. Examples of this situation in nature have been found beneath an Arctic ice island, in an Antarctic lake and possibly above the hot brine pools on the bottom of the Red Sea.

It is important to find an understanding of the generation, maintenance and decay of this microstructure. This will probably need the technically more difficult observations of the vertical shear of the currents.

The relation of the vertical fine structure and fluctuations observed in current meter records has not been fully explored. It is vital to an understanding of the general circulation of the ocean and the associated density structure.

Most of the vertical transfer of heat in the stably stratified ocean seems likely to occur through local overturning and billow instabilities associated with internal waves. Its relation to the horizontal transfer of momentum and the observed fluctuations in ocean currents is still to be elucidated. The

horizontal transfer properties of stably stratified rotating fluids like the atmosphere are closely related to the structure of the baroclinic eddies⁹. It may be that the oceanic equivalent of the atmospheric depression is crucial to an understanding of the ocean circulation. An international multiple ship expedition (MODE 1) is being planned with this in mind.

¹ Belderson, R. H., Kenyon, N. H., and Stride, A. H., *Mar. Geol.*, **12**, A11 (1972).

² Laughton, A. S., Whitmarsh, R. B., Rusby, J. S. M., Somers, M. L., Revie, J., McCartney, B. S., and Nafe, J. E., *Nature*, **237**, 217 (1972).

³ Gregg, M. C., and Cox, C. S., *Deep-Sea Res.*, **19**, 355 (1972).

⁴ Pingree, R. D., *Deep-Sea Res.*, **18**, 485 (1971).

⁵ Woods, J. D., and Wiley, R. C., *Deep-Sea Res.*, **19**, 87 (1972).

⁶ Thorpe, S. A., *Proc. 1972 Conf. IURCM Colloquium, San Diego, Calif.* (1972).

⁷ Tait, R. I., and Howe, M. R., *Nature*, **231**, 178 (1971).

⁸ Turner, J. S., and Stommel, H., *Proc. US Nat. Acad. Sci.*, **52**, 49 (1964).

⁹ Green, J. A. S., *Quart. J. Roy. Met. Soc.*, **96**, 157 (1970).

¹⁰ Rusby, J. S. M., *Int. Hydr. Rev.*, **47**, 25 (1970).

Drifting Through the Year

D. H. TARLING

School of Physics, University of Newcastle upon Tyne NE1 7RU

Many of the difficulties associated with attempts to apply plate tectonics to geophysical phenomena are a result of applying the basic tenets in too simple a manner.

DURING the past year or so some 100 km³ of new crustal rocks have been generated at the crests of the oceanic ridges, which must be comparable with the amount of newsprint devoted to the subject; like the ocean crust, most will be recycled after a comparatively brief appearance. It is difficult, under such circumstances, to select that fraction which will float to the top and be preserved and incorporated into our evolving concepts. On this basis, this article does not contain references. The most significant development, other than the appearance of textbooks based on concepts of plate tectonics rather than inserting the odd paragraph or two, is an increasing appreciation that many of the basic concepts of continental drift and plate tectonics are really only broad generalizations; there is increasing reappraisal of the reliability of the criteria in specific regions and a closer definition of areas where further testing is critical to the successful sorting out of the variety of hypotheses which have become incorporated under the blanket title of "plate tectonics". In general, however, the bandwagon continues to roll and the ramifications of the hypotheses have become increasingly significant throughout the Earth sciences.

This increased degree of healthy criticism is well illustrated by developments within the simplest of the continental drift arguments—the fitting of the continents. The original fit by Bullard *et al.* matched "old" parts of the continental shelves of the continents bordering the Atlantic by setting the computer to obtain the best fit, defined as being the minimum overlap and gap. Recent studies of the evolution of continental margins clearly show that even these old shelf edges are not reliable markers as the process of separation of the continents results in brittle fracture in the top 10 km or so of the continental crust, with ductile flow becoming increasingly important at depth. Such processes have led to the formation of marginal continental basins and have thereby modified the edges of the continents so that overlaps are more permissible and the best fit now tends towards zero gap—although obviously the true best fit is where geological structures can be

matched precisely. A similar realistic approach is also beginning to appear in the context of poles and angles of rotation. These are usually quoted to the first decimal place of a degree, yet a number of combinations of poles and angles will afford significant fits which may, in fact, satisfy other criteria better than the simple mathematical best fit. It has also become obvious that the average pole of rotation of an individual plate cannot be determined from a single transform fault as different faults may give significantly different poles for the same period of time; this reflects the fact that the plate edges at a fracture zone may not be as rigid as thought previously. The transform, in fact, constitutes a compromise between its previous evolution and subsequent plate motion. Nonetheless, the rigid concept remains paramount, although the significance of vertical movements, some of which may be unrelated to the horizontal motion of plates, is beginning to re-establish itself within the context of plate tectonics.

The recognition of key areas is well marked by the concentration of interest in such areas as the Bahamas—an area critical to the fit of North America and Africa and thus to the evolution of the Caribbean. There has been, however, surprising uniformity in the various assessments of the evolution of the Tasman Sea, with New Zealand fitting close to Australia/Antarctica until 80 m.y. ago after which New Zealand and the Lord Howe Rise separated until the opening of the ocean between Australia and Antarctica caused a northward component to appear some 50 m.y. ago. Similar uniformity also seems to exist in the interpretations of the evolution of the Palaeozoic Tasman geosyncline in South-east Australia. Such agreement does not extend into South-east Asia, an area that is both complex and little known from the point of view of plate tectonics, so that it is not surprising that the variety of hypotheses which have appeared in the past have continued to be supplemented by new hypotheses. The result has been a considerable increase in readily available information and it seems probable that more realistic models for the evolution of this area will evolve in the next few years as new concepts of island arc and marginal basin evolution, such as those of Karig, are supplemented by further data from geochemical studies and deep sea drilling. Nearer Britain this concentration on key areas is marked by several studies of the Rockall region and more detailed studies of the complex history of both the Norwegian and Labrador Seas—the latter assuming increasing significance as a potential site of mineralization and oil accumulation.

Mountains, Ocean Floors and Fossils

The relationship between plate motions and mountain building has turned out to be much closer than a simple compression between colliding plates or crumpling near a downgoing trench system. This has been demonstrated by the correlation by Coney of changes in direction of motion of the West Atlantic plate with the onset, and nature, of different orogenic events in the Western Cordillera of North America. Clearly such a detailed correlation not only emphasizes the very rigid nature of the plates but must lead to a greater understanding of the nature and changes in orogenic structures in other parts of the world—with the additional possibility of reversing the situation and determining changes in motions of plates from orogenic features—possibly for plates which no longer exist. Standard plate tectonic interpretations, without such detailed correlations, are becoming much more common for the classical orogenic areas, such as the Western Mediterranean, partially following the dating by the Deep Sea Drilling Project but chiefly returning to an examination of the development of such areas as the Tyrrhenian Sea and the Straits of Gibraltar in terms of important horizontal movements; the return is to a picture almost identical to that of Argand in the 1920s, except that the stable blocks are now called rigid microplates and these are now lost in the upper mantle instead of the lower crust. Clearly there is a bright future for such disposable small plates and these will almost certainly “explain” how the pre-Miocene rotation of the Iberian Peninsula, forming the Bay of Biscay, could occur without causing large compressive features in the Pyrenees. Although such reinterpretations are in progress on a small scale, it is, however, salutary to note that there remain many gross uncertainties, such as the past position of the continents for times when there is no record of oceanic magnetic anomalies available. This uncertainty was emphasized during the past year by the changes in the accepted position of Madagascar within the reconstructed Gondwanaland.

The position of Madagascar within any reconstruction of the southern continents is absolutely critical. For many years geophysicists have retained the du Toit model, which has South America and Africa united, as on all reconstructions, but Madagascar close to the Kenya coast, India against Madagascar on one side and Antarctica on the other, and Australia perched on Antarctica a little further round. During the past year, however, there has been increasing publication or repetition of geological and geophysical evidence against a northern origin for Madagascar and an increasing argument for a southern derivation, mostly from the Mozambique area. The evidence against a Kenyan origin appears to be definitive, so the south-westerly origin seems at least probable; this means that the entire arrangement of the continents forming Gondwanaland must be readjusted, and the subsequent history of breakup must be modified. Such rearrangements may not be serious in terms of past floral and faunal migration routes for, in spite of possible climatic difficulties, Antarctica still offers the principal connexion—although the possible inclusion of Southeast Asia is of great significance. This will sort itself out as it becomes clear where to look for the critical evidence, but it is symptomatic of the reliability of much of the available data that such gross positional changes can still be made and yet still be consistent with the evidence.

The search for old ocean floors continues with varying degrees of success. In particular, two separate slabs of cool lithosphere beneath the Western Cordillera of the United States have been recognized as remains of the Farallon plate. Such relict slabs are necessarily recently defunct, that is during the past few tens of m.y., if they are to be detected either by an associated heat flow anomaly or by a faster transmission of seismic waves through them. As the slabs gradually warm to the ambient temperatures of the upper mantle, the anomaly in travel times of 5 to 10% gradually becomes lost within the general level of lateral inhomogeneities within the upper mantle, of the order of 1%. Ancient “sinks” can only be detected from

the geological scars which they leave, either in the form of serpentinite remains of the oceanic lithosphere or geochemical markers, such as the K/Si ratio and lanthanide element distribution in andesitic volcanoes which can be directly related to the depth of the Benioff zone, while both were active. In general, most recent work seems to have been devoted to remapping outcrops of blueschists and so on, particularly in Canada and Alaska and along the line of the ancient Tethys between Europe and Africa; geochemical studies have chiefly been concerned with confirming the correlation in present-day seismic zones.

On the fossil front, the principal features are probably the increasing amount of information from Antarctica and the realization that what palaeontologists established last century has become accepted and validated by geophysicists who previously disproved it. This acceptance has led to a flurry of papers relating the dispersal and isolation of species to environmental changes as new oceans form or disappear, and it has become clear that several fossil organisms, such as plants and even birds, may be more diagnostic of past land connexions than has previously been thought. Conversely, an understanding of the cause of many morphological features, such as tree growth rings, is allowing a more realistic evaluation of such indicators of past climatic zoning. This understanding highlights many of the difficulties which have resulted from the grossly oversimplified use of the distribution of evaporites, tillites, coal and so on, for palaeolatitude indicators, and Earth scientists owe a great deal to those Devil's advocates, such as the Meyerhoffs, who continue to press for a more realistic assessment of the available evidence—even though their zeal may sometimes obscure their judgment!

Palaeomagnetism and Drilling

Significant advances have been made in palaeomagnetic studies of Precambrian rocks, particularly from Africa and Canada, and Irving's correlation between significant changes in the polar wandering and tectonic events in the evolution of the Canadian shield is vital, not only in terms of the history of the Canadian Shield but because it shows that those were continental motions in parts of the world well back into the Precambrian. Reappraisal of much of the Precambrian geology of Africa has, however, indicated little movement between the different cratonic blocks during various later “mobile” phases. Unfortunately the palaeomagnetic story for the Lower Palaeozoic remains obscure although there has been some improvement in the available data from Australia. But increasing doubt has been cast on the exactness of the axial geocentric dipole model for the geomagnetic field in the past. More data from Tertiary rocks seems to indicate that the average geomagnetic dipole, although axial, tends to be displaced northwards with the result that the palaeomagnetic determinations from any one region tend to lie on the far side of the present rotational pole—although oceanic magnetic anomalies would not permit such a displacement to be interpreted as a consequence of plate motions. If this situation continues to be verified, then it does not invalidate reconstructions based on matching polar wandering curves, but it does suggest that much more caution must be exercised in evaluating palaeolatitudes. This offset does not, however, appear to have been found in studies of deep sea sediments. Unfortunately, it still proves difficult to segregate true polar wandering from plate motions and a fuller comprehension of the mechanisms of plate motion may solve this apparent discrepancy.

Drilling

Deep sea drilling is still providing vital new evidence. The northward movement of the Pacific plate has been confirmed by the retrieval of sediments attributed to the equatorial sedimentation belt of 125 m.y. ago from the vicinity of the Aleutian Trench. The Wharton Basin of the East Indian Ocean is now known to be of Tertiary age, and not the oldest extant ocean

floor. The significance of such discoveries has yet to be assessed; in particular the Wharton Basin indicates a previous spreading from north to south and, therefore, a spreading ridge system which has disappeared beneath the Java Trench. The chief theoretical, rather than regional, discovery is, however, that the oceanic basins on the continental side of the island arc systems are of Tertiary age and therefore attributable to some secondary spreading system, similar to that of normal oceanic basins, although the rates are much slower. The only pessimistic note is that the length of oceanographic reports continues to be matched by the length of the list of authors' names. Does each author really contribute one sentence or is this merely to increase everybody's publication list? If so, then I want to add a full stop.

The emphasis in studies of oceanic magnetic anomalies seems to have switched to consideration of the two ends of the spectrum of polarity changes of the Earth's magnetic field during the past 150 m.y. or so. Detailed studies to delineate the series of reversals during the past 2 to 3 m.y. is leading to increased clarification of the number and duration of such brief events when the Earth's magnetic field changed polarity for only a few hundred thousand years or so. At the other end, more careful study is being made of the so-called magnetic quiet zone, for example, the 250 km wide area between the magnetically "noisy" record of the West Atlantic and the edge of the North American continental shelf. Patterns are beginning to appear within this zone which may be crucial in dating other linear patterns observed in such key areas as the Labrador Sea, the Weddell Sea and in the ocean to the south-west of Madagascar. More precise dating of these older polarity sequences may soon be forthcoming from studies of the remanent magnetization of continental rocks. Unfortunately, this seems to have got lost in semantic disputes as to the use of names (and if so, whose) or numbers for the different polarity changes. Surely the standard stratigraphic nomenclature is sufficiently complex without setting up a new system for polarity zones? This development is particularly to be deplored in a year which is otherwise marked by a welcome absence of new jargon.

Mineral Deposits

Economic considerations are beginning to get past the arm waving stage as it becomes clear that there are more fundamental considerations associated with continental drift than the simple matching of metallogenic zones on various reconstructions for different times. Naturally, such matching can be of considerable significance, but the concentration of ores is also clearly associated with the onset of separation—the high heat flows causing mobilization of ore minerals previously disseminated within the continental crust—as well as the standard association with granitic emplacement linked to the trench/island arc regions. Palaeolatitude considerations will probably become more pertinent, and there is a clear case for re-evaluating the possible palaeolatitudinal control of petroleum and gas deposits which has not yet been properly established. But the history of continental margin basins is clearly of direct relevance to most oil producing areas.

The discovery of ore deposits related to the opening and closing of past oceans will depend on plate tectonic mechanisms having operated in the past, so it is of great economic as well as theoretical importance that the activity of plate motions can be traced well into the Precambrian, particularly on ophiolite and palaeomagnetic evidence. The dating of the oldest continental rock from Godthaab, south-west Greenland, at 3,800 m.y. suggests that differentiation of granitic crust took place at a very early stage and would probably be associated with horizontal motions, although the change to "normal" plate tectonics probably took place when the Archaean granulite/greenstone association ceased to be dominant. There is increasing evidence for such plate motions, certainly back to 2,500 m.y. ago, and from theoretical considerations of the rapid formation of the Earth's history it seems that the present geodynamic system

would have been initiated shortly after this period of world wide heating. Further assessment must depend on developments in the understanding of the formation of the Earth and the mechanisms of continental drift.

What Progress?

In some ways there has been little advance in understanding the mechanisms of continental and plate motions during the past year. The two principal features have been the presentation of the "plume" theory in greater detail and the increased study of the role of phase transitions within geodynamic theories. The work of Ringwood is particularly important in that he presents the action of phase transitions, not for the first time, as the driving force on the lithospheric plates. The argument is that in a cool descending plate, the phase transitions are displaced upwards by some 30 to 100 km, so that the increase of 5 to 10% in density accompanying such transitions increases the density of the slab against that of the ambient mantle, causing the slab to descend and dragging the rest of the plate with it. (In this regard, it is highly significant that many seismologists are considering that deep and intermediate earthquakes may well be caused by phase transitions if they affect a volume of several km³. Such an interpretation is not new, but its wider acceptance is certainly important in any interpretation of the tectonics of a descending slab.) The chief difficulty with most phase transition arguments is that it is easy to see how such a mechanism may operate in one region, in this case at oceanic trenches, but the converse situation must operate in another region, such as the oceanic ridges, and any mechanism assisting descent in one place must restrict ascent in the other. This conceptual problem seems to have dogged most geodynamic considerations, although a similar blockage appears to arise when such systems are viewed against the way in which plates have evolved; for example, no explanation is usually provided of how a symmetrically spreading ridge commences beneath a continent, which it then separates along old lines of weakness, and subsequently migrates away from, say, Africa or Antarctica, yet still retains its symmetry irrespective of the presence or absence of a trench on the same plate. In this regard the plume theory, as developed by Jason Morgan from the initial concepts of Tuzo Wilson (as are so many plate tectonic ideas), has some attractions. The basic argument is that plumes of deep mantle material, some 150 km in diameter, rise through the entire mantle and spread out within the lithosphere. The plume ascent is marked by hot spots such as Hawaii, Iceland, Galapagos, and so on, and the plates, with accompanying trails of aseismic ridges, are carried away from the hot spots as the plume spread out. The concept is particularly useful in explaining how domes form on continents, the crests of which may crack to form, for instance, the East African rifts and the Rhine Valley graben, and clearly such cracks could then be extended to form a distinct line along which separation could take place. It is not clear, however, why such a mechanism should cause plates to move perpendicularly away from oceanic ridges rather than radiate from hot spots, and theoretical considerations of the behaviour of circulating material containing radioactive materials, which is therefore generating heat as it circulates, suggest that any such flow is concentrated into downward plumes, but has diffuse upward rises. Clearly the mechanism problem is still wide open and it is to be hoped that the Geodynamics Commission will at least define the boundary conditions within which any such mechanism must be confined.

All in all, it has been a stimulating year, with a plethora of ideas, some new, some revamped, but the principal feature has been a definition of the significant anomalies and difficulties which follow from an over simple application of the basic premises of plate tectonics. Clearly such reappraisal will lead to a much clearer understanding and less dogmatic approach to a subject which continues to revolutionize and galvanize the Earth sciences.

Changes in Mother–Infant Relationship after Separation in *Rhesus* Monkeys

ROBERT A. HINDE & LYND A. M. DAVIES

MRC Unit on the Development and Integration of Behaviour, Madingley, Cambridge

Infant distress after mother–infant separation depends greatly on the re-establishment of maternal care on reunion.

If separated from their mothers for a few days, infant *Rhesus*¹ monkeys and other macaques² show increased distress calling, and reduced locomotor activity and other symptoms of behavioural depression. The symptoms may persist for weeks after reunion with the mother, and infants which have had such a separation experience differ from non-separated controls in their responses to strange objects months and even years later³. The trauma of the separation experience itself may not be the only or even the major factor in producing these effects. The disturbance to the mother–infant relationship which results from the separation seems to be of considerable importance.

Mother–Infant Relationship after Reunion

In experiments involving removals of mothers from sixteen 21 to 32-week-old infants for six days⁴ and from five 30 to 42-week-old infants for thirteen days⁵, the mothers acceded only partially to the increased demands for nipple contact made after reunion. Over this age range non-separated infants spend progressively less time on their mothers. Separated infants spent more time on their mothers after reunion than before ($P < 0.01$, Wilcoxon 2-tailed test, for the six day group and 5/6, *n.s.*, for the thirteen day group, for the day of reunion; $P < 0.01$ and 5/6, *n.s.*, for the next week: in one of the 13 days group relations with the mother were not re-established for over a week), and this occurred in the absence of any significant reduction in the proportion of the infant's attempts to gain the nipple that were rejected by the mother⁶. Indeed, the proportion of nipple contacts terminated by the mother was markedly greater after reunion than before ($P < 0.02$ for both day of reunion and the next week for the 13 day group: adequate data were not available for most 6 day animals, but similar trends were apparent). In 13 day animals median increased from 0.48 pre-separation to 0.94 in second week after reunion, compared with a mean change of +0.01 ($N = 14$) in non-separated controls). By rejection of a high proportion of the infants' attempts to gain the nipple (Fig. 1) and termination of a high proportion of nipple contacts the mothers play a large part in determining the nature of the mother–infant interaction.

A similar conclusion is indicated by the proportion of time off their mothers that the infants spent out of arm's reach (> 60 cm). After reunion this was reduced in nearly all infants ($P < 0.01$ and 5/6, *n.s.*, in the two experiments for the day of reunion; $P < 0.01$ and 5/6, *n.s.*, for the next week). The infant's role in maintaining proximity to its mother can be assessed by the difference between the percentage of approaches due to movement by the infant and the percentage of leavings due to movement by the infant ($\%Ap - \%L$), where approaches and leavings occur when the infant's position changes with respect to a circle of 60 cm radius round the mother⁷. Positive values indicate that the infant is responsible for a higher proportion of approaches than it is of leavings, and thus is primarily responsible for maintaining proximity. After reunion this index tended to increase (8/10, *n.s.*, and 3/6, *n.s.*, on day of reunion; 12/16, $P < 0.02$ and 5/6, *n.s.*, in next week). Thus the decrease in time that the infants spent at a distance from their mothers occurred even though the mothers were playing a smaller role in maintaining mutual proximity.

We conclude that after temporary removal of the mothers, the mothers play a large part in limiting the closeness of the relationship after reunion.

Distress shown by the Infant after Reunion

The degree of distress shown by infants during separation and after reunion varies markedly between individuals. For the infants whose mothers were removed for 6 days, distress was measured by frequency of distress calling, the frequency with which they sat doing nothing, and the inverse of their locomotor activity. This distress index was found to be significantly correlated with rejection of attempts to gain the nipple and maintenance of proximity to the mother. Rank order correlation coefficients between pre-separation frequency of rejections and distress index for day of reunion, next week and days 13–27 after reunion respectively +0.18, *n.s.*; +0.59, $P < 0.05$; and +0.43, $P < 0.05$. Correlations between measure of

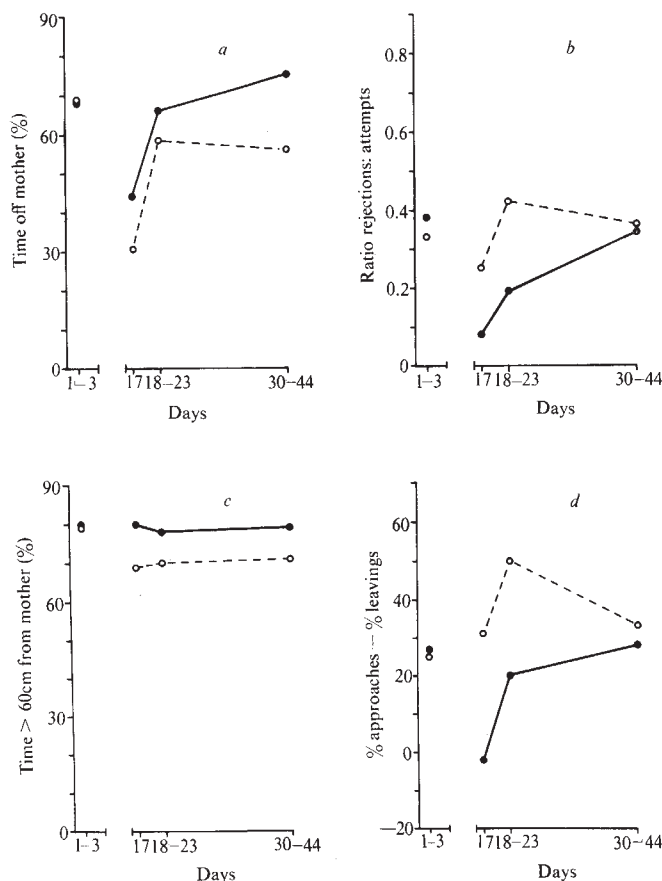


Fig. 1 Effects of 13 day separation on mother–infant relations. Separation was made on day 3. *a*, Number of half minutes in which infant was off mother as percentage of number observed. *b*, Ratio of number of times infants' attempts to gain nipple were rejected by mother to total number of attempts. *c*, Number of half minutes in which infant was > 60 cm from mother as percentage of number in which off her. *d*, Difference between percentage of approaches due to infant and percentage of leavings due to infant. ○---○, Mother removed; ●—●, infant removed.

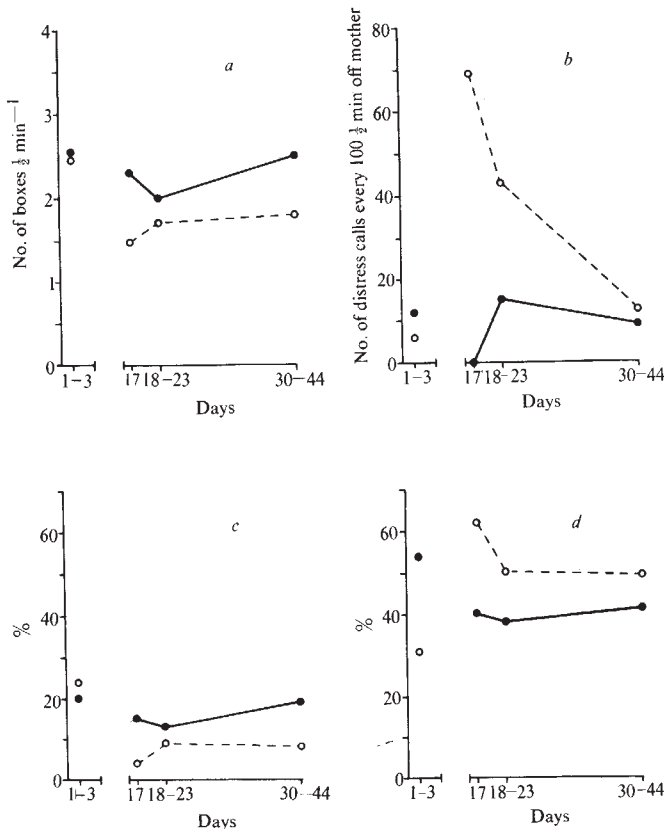


Fig. 2 Effects of 13 day separation from mother on locomotor behaviour of infant *Rhesus* monkeys when off mothers. ●—●, Median for infant-removed infants; ○---○, that for mother-removed infants. *a*, Separation was made on day 3, number of sections of cage entered per half minute off mother. *b*, Frequency of distress calls. *c*, Number of half minutes in which infant entered more than 5 sections of cage, as a percentage of number in which it entered more than one. *d*, Number of half minutes in which infant sat doing nothing as a percentage of number off mother.

infant's role in maintaining proximity and distress index respectively $+0.30$, *n.s.*; $+0.64$, $P < 0.01$; and 0.41 , *n.s.* The distress index immediately after reunion was more closely related to the pre-separation mother-infant relationship than post-separation, but by 2–4 weeks after separation the situation was reversed⁸. Although these correlations do not prove that individual differences in the mother-infant relationship determined differences in infant distress, this seems highly probable.

Infant-removed and Mother-removed Infants

In the previous experiments involving temporary removal of the mother, the infant remained in the same physical surroundings and with the same social companions. If the infant were removed from the mother, the infant should be more severely affected, for the trauma of the confinement in a strange situation would be added to that of separation. Data on 5 infants removed from their mothers for 13 days (infant-removed infants) were compared with those on 6 whose mothers were removed (mother-removed infants). In each case data were collected for 3 days before separation, mother and infant were separated on day 4 and reunited on day 17, and subsequent data were pooled for days 18, 19, 21 and 23 and for days 30, 37 and 44. Differences between the physical conditions of the cages in which these infant-removed infants were kept and the home pens in which the mother-removed infants remained prevented precise comparisons, but it did seem that the infant-removed infants were the more disturbed. After reunion, however, the reverse was the case⁹. The infant-removed infants showed more locomotor activity ($P = 0.04$, 0.03 and 0.01 for days 17, 18–23 and 30–44, respectively, Mann Whitney 2-tailed), were more active when active (*n.s.*, *n.s.* and $P = 0.01$), and gave

fewer distress calls ($P = 0.004$, 0.03 and *n.s.*), than infants whose mothers had been removed (Fig. 2). They also engaged in significantly more social play.

The small degree of distress shown by the infant-removed infants was accompanied by a smaller disruption of mother-infant relations (Fig. 1). The time spent off the mother, greater in the infant-removed infants than in the mother-removed ones (*n.s.*, *n.s.* and $P = 0.05$ on days 17, 18–23 and 30–44 respectively), was associated with a lesser relative frequency of rejections ($P < 0.06$, 0.06 , *n.s.*, Sign test). There was, however, no significant change in the proportion of nipple contacts made or broken by the infants. The improved relationship was thus primarily due to greater receptiveness by the mothers to the infants' demands for nipple contacts. By contrast, in the mother-removed group the relative frequency of rejections and the infant's role in making nipple contacts changed little, but there was a marked decrease in the proportion of nipple contacts broken on the infant's initiative ($P < 0.02$, 0.02 and 0.02). The mothers not becoming more receptive played a greater role in terminating nipple contacts.

Similar considerations apply to the relative roles of mother and infant in maintaining proximity. The infant-removed infants spent a slightly higher proportion of their time off their mothers at a distance from her than did the mother-removed infants (*n.s.*, *n.s.* and $P = 0.01$). The infants' role in maintaining proximity tended to be greater in the mother-removed infants (*n.s.*) than before separation, and less ($P = 0.06$, *n.s.*, *n.s.*) in the infant-removed infants, though the inter-group differences did not reach significance.

Thus the smaller distress shown by the infant-removed infants after reunion was associated with a smoother re-establishment of mother-infant relations, and the latter was due to a readier acceptance by the mothers of the demands for nipple contact and proximity made by the infants after reunion. Presumably the mothers, whose role in determining the nature of the mother-infant relationship after separation has already been emphasized, suffer less stress if their infants are merely removed than if they are both separated from their infants and placed in a strange environment. These findings support the view that the disturbance to the mother-infant relationship caused by the separation episode is an important factor in the distress shown by the infant in the post-separation period.

Two further points must be made. First, this suggestion does not imply that separation *per se* has no subsequent effects. The infants show marked distress during separation, and it would be surprising if this had no later consequences. Second, our opinion is that there are no human parallels, though no data exist. Disturbance to the relationship is found in both species, however. While the view that separation from the mother may have damaging effects on the development of the human child¹⁰ is well substantiated¹¹, there is growing evidence that the crucial factor in separation is the deprivation of maternal care which may, but need not, accompany it¹².

This work was carried out in collaboration with the late Yvette Spencer-Booth and was supported by the Medical Research Council and the Royal Society.

Received April 27, 1972.

¹ Hinde, R. A., and Spencer-Booth, Y., *Science*, **173**, 111 (1971).

² Rosenblum, L. A., in *The Origins of Human Social Relations* (edit. by Schaffer, H. R.) (Academic Press, London and New York, 1971).

³ Spencer-Booth, Y., and Hinde, R. A., *J. Child Psychol. Psychiat.*, **12**, 157 (1971).

⁴ Spencer-Booth, Y., and Hinde, R. A., *Anim. Behav.*, **19**, 174 (1971).

⁵ Spencer-Booth, Y., and Hinde, R. A., *Anim. Behav.*, **19**, 595 (1971).

⁶ Hinde, R. A., *Ann. New York Acad. Sci.*, **159**, 651 (1969).

⁷ Hinde, R. A., and Atkinson, S., *Anim. Behav.*, **18**, 169 (1970).

⁸ Hinde, R. A., and Spencer-Booth, Y., *J. Child Psychol. Psychiat.*, **11**, 159 (1970).

⁹ Hinde, R. A., and Davies, L., *J. Child Psychol. Psychiat.* (in the press).

¹⁰ Bowlby, J., *J. Ment. Sci.*, **99**, 265 (1953).

¹¹ Ainsworth, M. D., in *Deprivation of Maternal Care* (WHO, Geneva, 1962).

¹² Howells, J. G., *Acta Paedopsychiatrica*, **37**, 3 (1970).

LETTERS TO NATURE

PHYSICAL SCIENCES

Small "Non-velocity" Redshifts in Galaxies

THE question of whether or not the small excess redshifts observed in some companion galaxies^{1,2} have a non-velocity origin has recently been raised. The de Vaucouleurs³ have pointed out that small, systematic errors in the redshifts of E and SO galaxies might arise because the lines in the low dispersion spectra on which they are measured do not have well-determined effective wavelengths. The data reported here support this possibility. In the three cases I present, systematic differences arise between velocities derived from measurements of CA II, H($\lambda 3968.5$) and K($\lambda 3933.7$) carried out "classically" (with a visually positioned, two coordinate measuring engine) and velocities derived from either "classical" measurements of other spectral lines in the same object or measurements of the same H and K lines using numerical techniques.

The strongest evidence for such systematic velocity differences comes from a measurement technique which finds the rotation curve for a galaxy by cross-correlating segments of its spectrum with that of a star of similar spectral type. For this method to work the star and galaxy spectra must be taken with the same instrument under similar conditions, measured with a digital microdensitometer, converted to relative intensities and the stellar spectrum numerically "broadened"⁴. The results of such measurements for the linear part of the rotation curve of NGC 5866 (SO spectral type K) are plotted with open circles in Fig. 1. The comparison star used for these was η Draconis (G8 III). The measurements are all for a single galaxy plate, a 130 Å/mm, image tube spectrum taken at Kitt Peak, whose zero point was determined by measuring night sky lines.

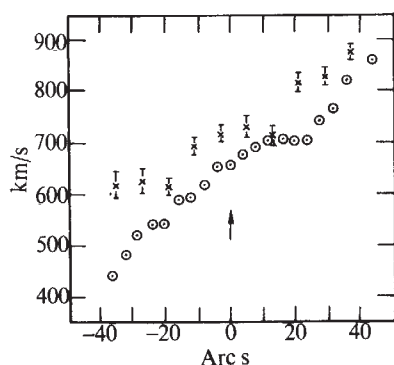


Fig. 1 Central rotation curve for NGC 5866. Points with error bars (x) are from "classical" measurements, those with open circles (○) from cross-correlation with the spectrum of η Draconis. Size of bars and circles is 2σ . The arrow indicates the nucleus.

Fig. 2 shows the cross-correlation calculated for three limited spectral regions of the galaxy's nucleus and suitably normalized. There is good agreement between the redshift determined from the region of H and K, where the continuum slopes steeply, and that found using the region 4000–4200 Å, which is relatively flat and void of prominent lines. The spectral regions 3800–

4000 Å, 3800–4200 Å, and 4000–4200 Å, give nuclear redshifts of 642, 664, and 649 km/s, respectively.

The points plotted with error bars in Fig. 1 were obtained by measuring H and K in the "classical" way: by setting on the absorption lines visually with a two-coordinate measuring engine. A smooth curve through these points gives a nuclear velocity of 725 ± 30 km/s.

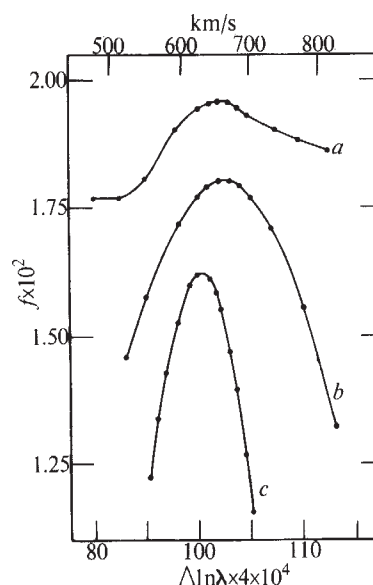


Fig. 2 Cross-correlation between spectra of η Draconis and of the nucleus of NGC 5866. Zero point for the ordinate is arbitrary. a, 4000–4200 Å; b, 3800–4200 Å; c, 3800–4000 Å.

Because the cross-correlation measurements for the H and K lines (3800–4000 Å) agree with those for the "flat" region (4000–4200 Å), the difference between the "classical" measurements and the "impersonal" computer calculation cannot be intrinsic to the CA II lines themselves. Moreover, the only possible source of systematic error between the two curves is the velocity of η Draconis relative to its comparison spectrum, which provides the zero point for the cross-correlation measurements. This velocity, however, is well determined (with a standard error of 13 km/s) and any error it introduces should be well within the observed difference of 73 km/s. Thus, the systematic shift between the "classically" measured points and those measured with the cross-correlation technique must arise from the different method of measurement. In short, measurements of Ca H and K absorption lines carried out "classically" require a different effective wavelength than those made using a template constructed from a stellar spectrum which closely mimics that of the galaxy.

The second line of evidence for a systematic shift comes from determining the centroids of the galaxy H and K lines numerically. This procedure gives a recessional velocity of 635 km/s when the slope of the continuum is left in, and 690 km/s when it is removed numerically. The latter value is in good agreement with a similar measurement by de Vaucouleurs⁵ of 692 km/s and is clearly an upper limit to the redshift. Both of these calculations lead to curves whose shapes are

identical to that defined by the circles in Fig. 1. In this case, since both the "classical" measurements and the profile measurements were done on the same plate, there is no possibility of a zero point difference.

Finally, earlier "classical" measurements on NGC 2903 (my unpublished data) and NGC 5194⁶ have shown a systematic difference between velocities from hydrogen lines (H β emission and H δ absorption) and those from Ca K absorption lines. These shifts, again, show Ca K velocities to be 40–60 km/s larger than those for H δ absorption or H β emission. Because all three of these lines are measured on the same plate simultaneously, there is no way in which a systematic zero point error could arise.

All these data suggest that velocities obtained from the H and K absorption lines of Ca II by "classical" methods give redshifts that are systematically too high by 40–70 km/s. This shift may arise from a tendency of the measurer to set the cross hairs on the denser, or redward edge of the sloping continuum in the region of these lines. Because redshifts for spiral galaxies are usually measured from emission lines while those for E and SO galaxies come principally from the absorption lines of Ca H and K and the G band⁷ one might well expect to find a systematic velocity difference between these two classes of objects. The possibility that the "classically" measured redshifts for E and SO galaxies are systematically too high could most easily be investigated by means of a device, such as that described by Lowrance *et al.*⁸ which records galaxy spectra directly in intensity, thus facilitating the type of analyses described here. An investigation of this type on NGC 224, 3031, 5128 and their companions would either support or disprove the small redshift discrepancies within these groups noted by Arp¹. But until such observations are conducted the evidence presented here would seem to suggest that any small "non-velocity" redshifts between spirals and SOs should be attributed to nothing more exotic than errors of measurement.

The microdensitometry and computing for this work were done at the Goddard Institute for Space Studies, NYC. Partial support was provided by a NSF grant.

SUSAN M. SIMKIN

Department of Astronomy,
Columbia University,
New York, NY,
and Kitt Peak National Observatory

Received June 22, 1972.

¹ Arp, H., *Nature*, **225**, 1033 (1970).

² Jaakkola, T., *Nature*, **234**, 534 (1971).

³ de Vaucouleurs, G., and de Vaucouleurs, A., *Nature*, **236**, 166 (1972).

⁴ Minkowski, R., *Problems of Extragalactic Research* (edit. by McVittie, G.), 112 (Macmillan, New York, 1962).

⁵ de Vaucouleurs, G., *IAU Symposium No. 30* (edit. by Batten, A., and Heard, J.), 16 (Academic Press, New York, 1967).

⁶ Simkin, S., *Bull. Amer. Astron. Soc.*, **1**, 362 (1969).

⁷ Humason, M., Mayall, N., and Sandage, A., *Astron. J.*, **61**, 97 (1956).

⁸ Lowrance, J., Morton, D., Zucchini, P., Oke, J., and Schmidt, M., *Astrophys. J.*, **171**, 233 (1972).

Cosmic Abundance of Iron and Nature of Primitive Material in Meteorites

A TENET of cosmochemistry is the identification of bulk analyses of Type I carbonaceous meteorites with abundances, of the condensable elements, in the primordial solar system^{1,2}. Such analyses are basic to most tabulations of cosmic abundances^{3–5}. A Type I meteorite is complex, mineralogically and petrologically^{6–9}, however, and there are grounds for questioning the extent to which it has remained a closed system since separation from the solar gas^{6,10}. I have suggested that the apparently primitive chemistry of such a

meteorite may be the result of a major component dominating the analysis and that this component might be unaltered solar system condensate¹¹.

In attempting to define the "primitive essence" of a Type I carbonaceous meteorite, some mineral constituents may be eliminated because they are demonstrably either mechanical additions (olivine and pyroxene^{6,8}) or late deposits (sulphates and carbonates^{6,12}). Of the remaining major constituents, no adequate genesis for the pyrrhotite has yet been advanced. Fortunately its abundance is quite low, about 2% by weight⁶, so that ignoring it has a small effect on the iron abundance. The main problem has to do with inter-relations between magnetite, which comprises 10–15% by weight of the meteorite^{13,14} and phyllosilicates, which occur as sub-microscopic, chemically heterogeneous grains making up about 65% by weight of the meteorite¹¹.

On the one hand, the phyllosilicate-magnetite association is seen as the result of hydration of primordial condensate by equilibration with the solar gas below about 315 K (ref. 15). This could have occurred before, during or after accretion, with a result resembling terrestrial serpentine-magnetite associations consequent upon aqueous alteration. On the other hand I suggest that phyllosilicates, representing unaltered material, and magnetite, representing altered material, were mixed together at low temperature, presumably on the meteorite parent body, so that they never equilibrated with each other. A further complication is that some phyllosilicates may be relatively late alteration products¹⁶. I shall return to this point, after considering arguments for the mixture hypothesis.

First, aqueous alteration of an iron-bearing silicate on Earth usually results in transfer of much of the iron into oxide or hydroxide phases¹⁷. The mean Fe/Mg ratio of phyllosilicates from a Type I meteorite, however, is about 1:2 (ref. 11) and the extreme heterogeneity observed makes it unlikely that this composition was fixed stoichiometrically. Possibly physico-chemical conditions were such as to favour retention of so much iron in phyllosilicates, but this point should be demonstrated rather than assumed.

Second, Type I phyllosilicates contain about 1% by weight of nickel^{6,11}. If they were formed by hydration of high-temperature silicates, such as olivine or pyroxene, some of this parent material should have survived, yet no other meteoritic silicate contains so much nickel¹⁸. Similarly, so high an Fe/Mg ratio is seldom found in meteoritic olivine or pyroxene¹⁹.

Third, silicate fractions from a Type I meteorite contain trapped rare gases with elemental and isotopic abundances of planetary type, whereas magnetite fractions are characterized by rare gases of solar type¹⁴, apparently implanted by a solar wind similar to that at present. This suggests that silicate and magnetite acquired their gases in different environments and were subsequently mixed at temperatures too low to permit equilibration.

Fourth, phyllosilicate-magnetite associations occur not only in Type I meteorites but also in the matrix of Type II carbonaceous meteorites. This matrix material closely resembles a Type I meteorite both elementally¹⁵ and isotopically²⁰. The highly characteristic morphologies exhibited by Type I magnetite, however, do not occur in Type II meteorites²¹, suggesting that only the silicate material is common to both and that the magnetites were late additions to the systems.

Fifth, the mean phyllosilicate composition matches the solar abundances, for major elements, at least as well as do bulk meteorite values (ref. 11 and Fig. 1). Although uncertainties in the solar data vitiate much of the usefulness of this approach at present (uncertainty bars in Fig. 1), the quality of the data must improve so that a real comparison should become possible. Meanwhile I conclude that existing data are compatible with the idea that the "primitive essence" of a Type I carbonaceous meteorite consists of its phyllosilicate population or a major part of it.

This hypothesis, if correct, leads to a new value for the cosmic abundance of iron of $5.3 \pm 0.6 \times 10^5$ (relative to Si = 10^6).

This value is sensibly different from the commonly used bulk value of $9.0 \pm 0.5 \times 10^5$, though both lie within the uncertainty limits of the solar spectroscopic data (Fig. 1). It is unlikely that abundance values for other elements would be greatly affected by this hypothesis. Phyllosilicate analyses for magnesium are of relatively poor quality and the 15% reduction in abundance suggested by Fig. 1 may not be real. Other elements determined so far, aluminium and nickel, show no significant change, reflecting, in part, the exceptional purity of Type I magnetite⁶.

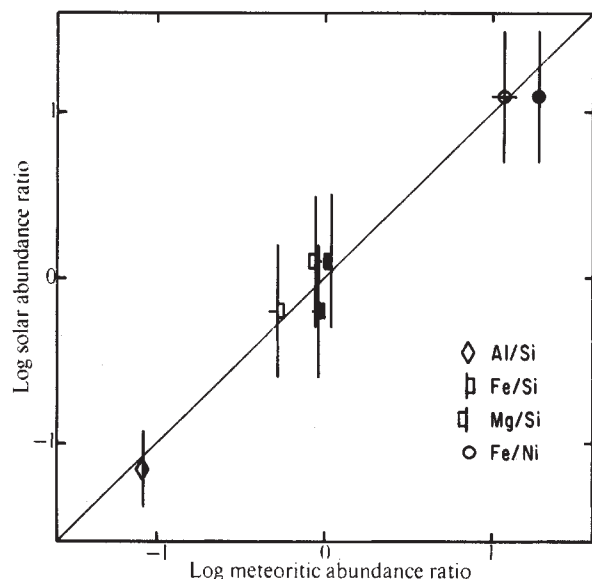


Fig. 1 Solar abundance data (based on refs. 26–28 and private communication from Aller and Ross) compared with bulk analyses of Type I carbonaceous meteorites (solid symbols, based on refs. 29–32) and microprobe analyses of Type I phyllosilicates (open symbols, based on refs. 6 and 11).

The idea that this phyllosilicate material is unaltered condensate from the early solar system^{11,22} is consistent with the scanty data from laboratory simulations²³. Black, however, has suggested that Type I carbonaceous meteorites contain an unaltered pre-solar system component²⁴, thus providing an alternative interpretation which fits the data at least equally well. Survival of some pre-existing material is also suggested by Cameron's latest evolutionary model of the solar system²⁵.

The hypothesis that Type I phyllosilicates may be taken as the building blocks of the solar system can be tested experimentally. Improved astronomical and laboratory determinations should permit solar abundance data to be critically compared with analyses both of Type I meteorites and, by contrast, of phyllosilicate separates from them. The mineral separation techniques should take account of the possibility that some phyllosilicates may be altered material¹⁶.

This research was partially supported by the US Air Force and by NASA, NRC, and the Smithsonian Research Foundation. I thank the British Museum and the Smithsonian Institution for the meteorite samples.

J. KERRIDGE

Institute of Geophysics
and Department of Chemistry,
University of California,
Los Angeles, California 90024

Received July 3; revised July 17, 1972.

¹ Urey, H. C., in *Nuclear Processes in Geological Settings*, 49 (NAS/NRC, Washington, 1953).

² Anders, E., *Geochim. Cosmochim. Acta*, **35**, 516 (1971).

³ Suess, H. E., and Urey, H. C., *Rev. Mod. Phys.*, **28**, 53 (1956).

- ⁴ Cameron, A. G. W., in *Origin and Distribution of the Elements* (edit. by Ahrens, L. H.), 125 (Pergamon, Oxford, 1968).
- ⁵ Goles, G. G., in *The Handbook of Geochemistry*, Part I, 123 (Springer, Berlin, 1969).
- ⁶ Boström, K., and Fredriksson, K., *Smithsonian Misc. Coll.*, **151**, 3 (1966).
- ⁷ Orcel, J., and Alpern, B., *C.R. Acad. Sci. Paris*, **D262**, 1393 (1966).
- ⁸ Reid, A. M., Bass, M. N., Fujita, H., Kerridge, J. F., and Fredriksson, K., *Geochim. Cosmochim. Acta*, **34**, 1253 (1970).
- ⁹ Bass, M. N., *Meteoritics*, **5**, 180 (1970).
- ¹⁰ Urey, H. C., *Quart. J. Roy. Astron. Soc.*, **8**, 23 (1967).
- ¹¹ Kerridge, J. F., *Nature Physical Science*, **230**, 66 (1971).
- ¹² DuFresne, E. R., and Anders, E., *Geochim. Cosmochim. Acta*, **26**, 1085 (1962).
- ¹³ Kerridge, J. F., and Chatterji, S., *Nature*, **220**, 775 (1968).
- ¹⁴ Jeffery, P. M., and Anders, E., *Geochim. Cosmochim. Acta*, **34**, 1175 (1970).
- ¹⁵ Larimer, J. W., and Anders, E., *Geochim. Cosmochim. Acta*, **31**, 1239 (1967).
- ¹⁶ Bass, M. N., *Geochim. Cosmochim. Acta*, **35**, 139 (1971).
- ¹⁷ Deer, W. A., Howie, R. A., and Zussman, J., *Rock Forming Minerals*, 3 (Wiley, New York, 1962).
- ¹⁸ Fredriksson, K., and Keil, K., *Meteoritics*, **2**, 201 (1964).
- ¹⁹ Van Schmus, W. R., *Earth Sci. Rev.*, **5**, 145 (1969).
- ²⁰ Taylor, H. P., Duke, M. B., Silver, L. T., and Epstein, S., *Geochim. Cosmochim. Acta*, **29**, 489 (1965).
- ²¹ Jedwab, J., in *Origin and Distribution of the Elements* (edit. by Ahrens, L. H.), 467 (Pergamon, Oxford, 1968).
- ²² Arrhenius, G., and Alfvén, H., *Earth Planet. Sci. Lett.*, **10**, 253 (1971).
- ²³ Meyer, C., *Geochim. Cosmochim. Acta*, **35**, 551 (1971).
- ²⁴ Black, D. C., *Geochim. Cosmochim. Acta*, **36**, 377 (1972).
- ²⁵ Cameron, A. G. W., in *CNRS Colloquium, Nice* (1972).
- ²⁶ Huber, M. C. E., and Parkinson, W. H., *Astrophys. J.*, **172**, 229 (1972).
- ²⁷ O'Mara, B. J., thesis, University of California (1967).
- ²⁸ Garz, T., Kock, M., Richter, J., Baschek, B., Holweger, H., and Unsöld, A., *Nature*, **223**, 1254 (1969).
- ²⁹ Mason, B., *Space Sci. Rev.*, **1**, 621 (1963).
- ³⁰ Greenland, L., and Lovering, J. F., *Geochim. Cosmochim. Acta*, **29**, 821 (1965).
- ³¹ Nichiporuk, W., Chodos, A., Helin, E., and Brown, H., *Geochim. Cosmochim. Acta*, **31**, 1911 (1967).
- ³² Ahrens, L. H., Von Michaelis, H., and Fesq, H. W., *Earth Planet. Sci. Lett.*, **6**, 285 (1969).

BIOLOGICAL SCIENCES

Jump of the Oriental Rat Flea *Xenopsylla cheopis* (Roths.)

THE spectacular and bewilderingly rapid jump of the flea has always excited the curiosity and wonder of naturalists. In an attempt to explain the mechanism of the jump we have re-examined the morphology of *Xenopsylla cheopis*, particularly the exoskeleton and musculature, and have combined this detailed study with a series of high speed cine films.

The required energy is generated and stored by the compression of the pleural arch (which contains resilin)^{1,2} and the bending of the pleural and coxal walls by muscular contraction, and is released by cuticular click mechanisms. The force thus liberated travels through a series of roughly vertical ridges ("lines of force") from the notum via, the trochanter and tibia of the third pair of legs to the substratum (Fig. 1). The claws and tarsi of these legs do not participate actively in the jump.

Fig. 2 shows sequential frames from a high speed cine film of a flea (*X. cheopis*) jumping. The claws and five segments of the tarsi can be seen extending over the edge of the narrow jumping platform. Frame by frame analyses of the position of the flea, relative to the substratum, together with a knowledge of the framing rate (3,500 frames s⁻¹) allows calculation of the flea's acceleration, which is proportional to the force generated by the flea (Fig. 3). The peak acceleration is 1,350 m s⁻² (roughly equivalent to 140g).

Although fleas, like the majority of mammalian and avian ectoparasites, have lost the ability to fly, they have retained

and modified several important features associated with a winged mode of locomotion and incorporated them into the jump mechanism. The resilin-containing wing hinge ligament is present in the form of the "pleural arch", which, owing to the lateral flattening of the flea's body, has come to occupy a lateral rather than a dorsal position³ and consists of a helmet-shaped mass of resilin capping the pleural ridge. The muscles involved in the arching (bending) of the pleural walls and the coxal walls, the compression of the resilin, and "locking" the pleural ridges into the rim of the coxae, include direct and indirect flight muscles, the trochanteral depressor (starter muscle) and the epipleural muscles (basalars and subalars, possibly combined with dorsoventral muscles).

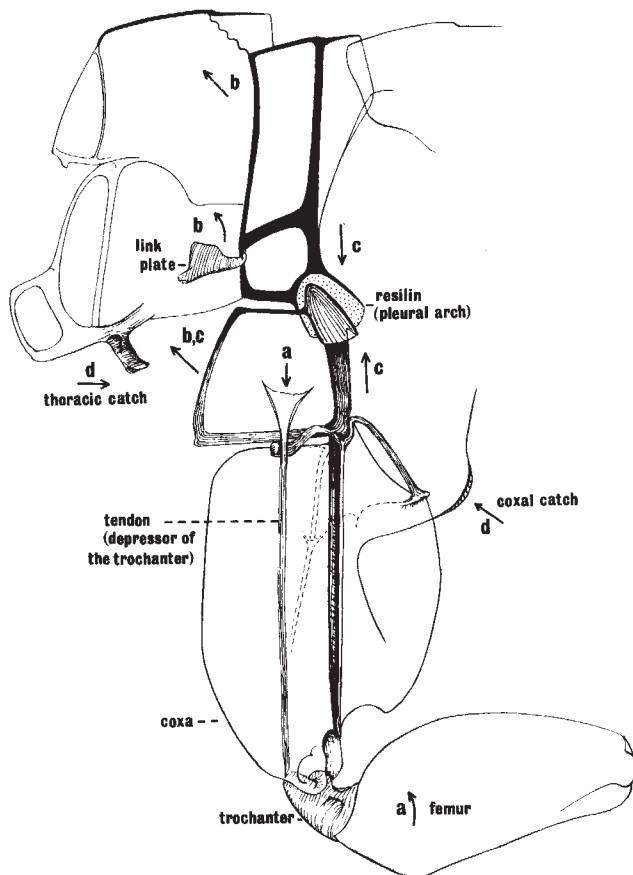


Fig. 1 Drawing showing the metathorax and third (jumping) leg in the relaxed position. The letters indicate the sequence of movements in preparation for a jump. *a*, The raising of the femur pulling the tendon of the trochanteral depressor downwards. *b*, Rotation of link plate raising the third segment which brings it into line with the second segment. *c*, Compression of the resilin (and arching of the pleural walls). *d*, Engagement of thoracic catch in membrane of metasternum, and coxal/abdominal catch.

When the flea prepares to jump it crouches down (Fig. 4) and lowers the head towards the substratum (Fig. 5), thus arching the back and contracting the body. The crouch is the result of elevating the femurs, rotating the trochanteral hooks into the sockets (concavities) in the distal end of the coxa, and bringing the trochanter and tibia of the third pair of legs into contact with the substratum. The first and second pair of legs (in this species) act chiefly as supports. This action involves a series of muscular contractions (of the longitudinal and oblique muscles of the thorax) which results in the upward rotation of the thoracic link-plates, clamping the three thoracic segments into position (Fig. 1) and engaging the cuticular catches (Fig. 6). Thus the thoracic segments are linked together ventrally and the coxa is fixed to the abdomen. The

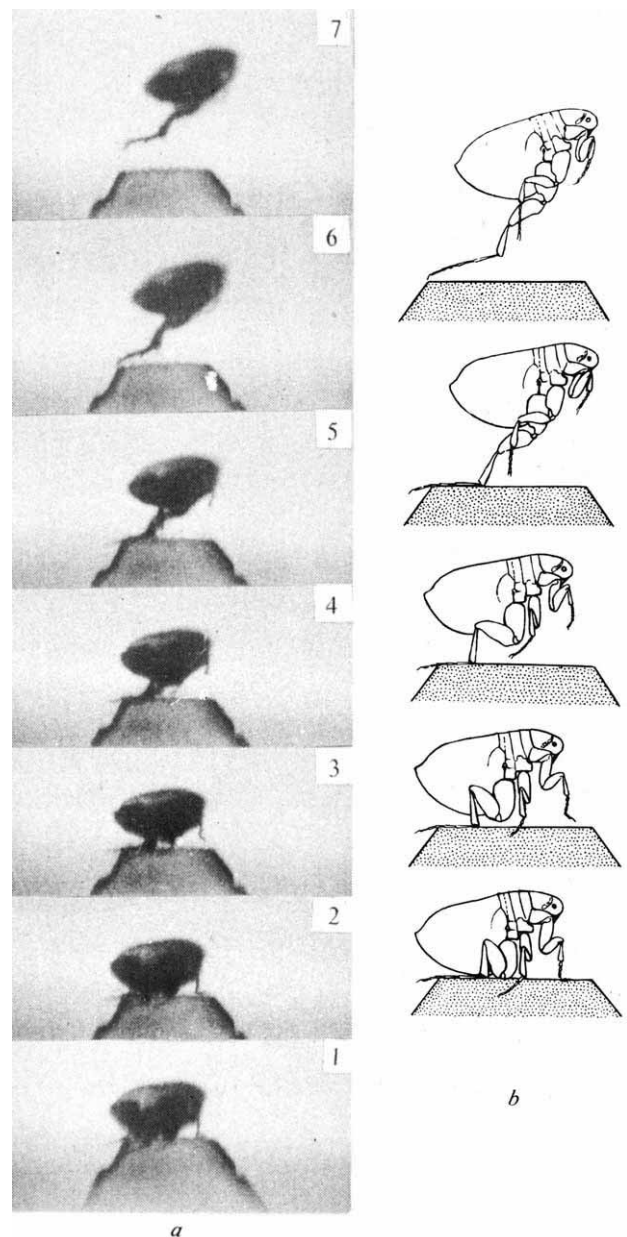


Fig. 2 *a*, Flea jumping (time between sequential frames 0.3 ms). *b*, Outline drawings made from frames 2-6.

resilin in the pleural arch is compressed by contraction of the trochanteral depressor and the epipleural muscles, and held thus by the thoracic catches. These are engaged by contraction of the ventral longitudinal muscles which also closes the gap between the pleural ridge and the outer rim of the coxa, thus making the line of force continuous. In this arched position the flea is "cocked" for take-off, and the muscles which compress the resilin can be relaxed.

The jump is initiated by the relaxation of certain muscles, the first being the levator of the coxa, allowing the femur to descend. Simultaneously the muscles holding the catches in place relax, releasing the energy stored in the resilin and the arched pleural and coxal walls. The coxal ridges straighten and the funnel-shaped upper portion of the tendon of the trochanteral depressor becomes wedged in its socket (Fig. 6) and braced against the metafurca. The trochanter is thus pressed downwards against the substratum and the force generated by the expansion of the resilin and straightening of the pleural walls is transmitted down the "line of force" formed by the heavily sclerotized vertical ridges of the notum, pleurum and coxa. This provides the initial acceleration. After the

trochanter leaves the ground the thrust is continued as the leg straightens and the descending femur exerts force through the tibia. The leverage for this movement is supplied by the point of attachment of the tendon of the trochanteral depressor, acting against the coxa-trochanteral hinge located at the base of the line of force. As in insect flight systems^{4,5} elasticity in both resilin and ordinary cuticle ensures very rapid movement as well as simplifying coordination. Synchrony of fast-moving parts is achieved with greater effect by the release of mechanical click mechanisms than by nervous coordination.

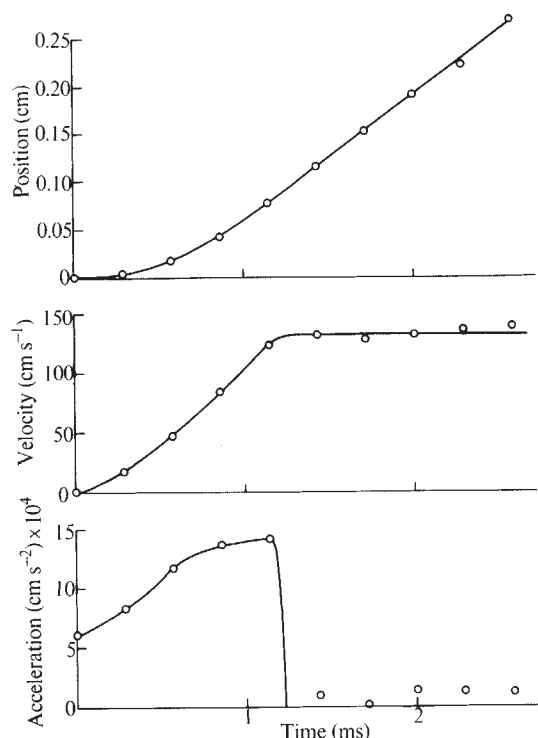


Fig. 3 Measured position, velocity and acceleration of a flea during its jump.



Fig. 4 Flea climbing on to pedestal.



Fig. 5 Flea preparing to jump.

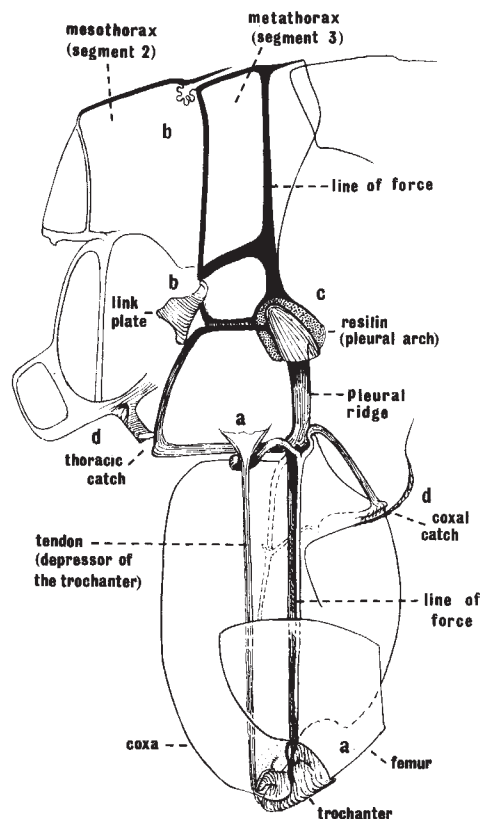


Fig. 6 Drawing, as Fig. 1, showing the different sclerites in position when the flea is ready to jump. *a*, Femur raised and tendon of trochanteral depressor held in the funnel-shaped socket. *b*, Link plate fully elevated. *c*, Resilin compressed (and arching of pleural walls). *d*, Thoracic and coxal catches engaged.

The photographs (Fig. 3) show clearly that the flea's claws and the five long tarsal segments of the back leg do not straighten, and remain parallel with the substratum after take-off. Frequently a pair of legs is flicked upwards and carried above instead of below the body⁵, giving the flea a better chance of attachment if it lands at an awkward angle on the host.

X. cheopis can still leap if the entire back leg is cut off, but the jump is then very feeble. The structure of the musculature and exoskeleton of the second leg suggests that the mechanism is essentially the same as that described for the back leg; however, in this case two of the catch mechanisms, the pleural arch and the resilin, are lacking and the pleural rod is not attached to the pleural wall. Thus no "line of force" exists.

MIRIAM ROTHSCHILD

Ashton, Peterborough

Y. SCHLEIN

Department of Protozoology,
The Hebrew University, Jerusalem

K. PARKER

Physiological Flow Studies Unit,
Imperial College, London SW7

S. STERNBERG

Department of Entomology,
Israel Institute for Biological Research,
Ness-Ziona, Israel

Received July 5, 1972.

¹ Bennet-Clarke, H. C., and Lucey, E. C. A., *J. Exp. Biol.*, **47**, 59 (1967).

² Neville, C., and Rothschild, M., *Proc. R. Ent. Soc. Lond. (C)*, **32** (3), 9 (1967).

³ Rothschild, M., *Proc. Brit. Nat. Hist. Soc.*, **1**, 1 (1969).

⁴ Pringle, J. W. S., *Insect Flight* (Cambridge University Press, 1957).

⁵ Andersen, S. O., and Weis-Fogh, T., in *Advances in Insect Physiology* (edit. by Beament, J. W. L., Treherne, J. E., and Wigglesworth, V. B.), **2** (Academic Press, London and New York, 1964).

⁶ Rothschild, M., *Scient. Amer.*, **213**, 44 (1965).

How Bees Navigate

WORKER bees communicate information about food sources by means of a dance in which the direction is given by the angle which the dancing bee makes with the vertical¹⁻³. The bees must use this angle relative to some reference outside the hive. They do this by orienting to the polarization of the sky^{1,3}.

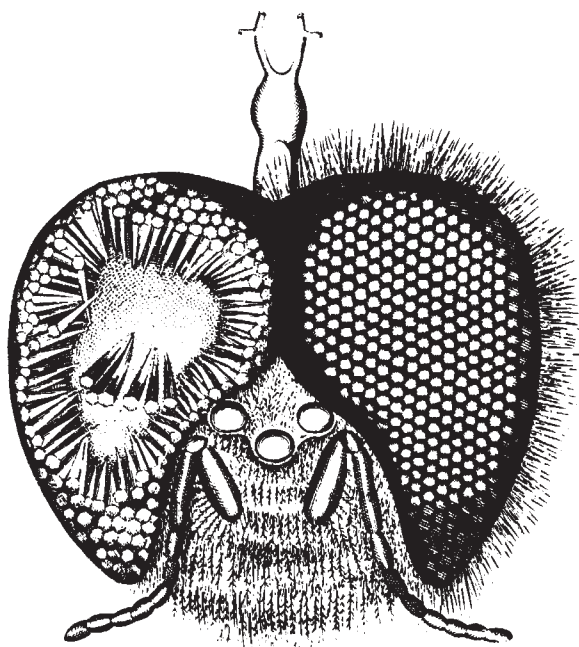


Fig. 1 Head of the honey bee showing the ommatidia of the compound eyes according to Swammerdam, *Bijbel der Natuure*, Leyden 1737.

The bee has a compound eye made up of thousands of light receptors known as ommatidia (little eyes) illustrated in Fig. 1. Recent morphological⁴ and optical⁵ studies provide an excellent description of the worker bee ommatidium, schematically illustrated in Fig. 2. Light enters through the cuticular lens (CL) and passes within the crystalline cone (CC) to the rhabdom (Rh) where the photosensitive pigment is located. The rhabdom, illustrated by Fig 2b, is approximately 4 μm in diameter, 350 μm in length, and has an index of refraction slightly larger than its surround. The photopigment is localized in tightly packed microvilli, represented by parallel dark lines in Fig. 2b. Each quadrant is formed by two reticular cells that sense the absorbed light. Because of the orientation of photopigment molecules in microvilli, the light absorption is sensitive to polarization. More light is absorbed if the electric vector is parallel to the microvilli than if it is perpendicular to them^{1,3}.

Here we show that, in addition to the microvilli, the physical structure of the worker bee rhabdom greatly enhances its polarization sensitivity by selectively detecting certain electromagnetic modes that are transmitted along it. Our theory accurately predicts the spacing between the ommatidia responsible for direction finding.

Light is transmitted along the rhabdom in the form of

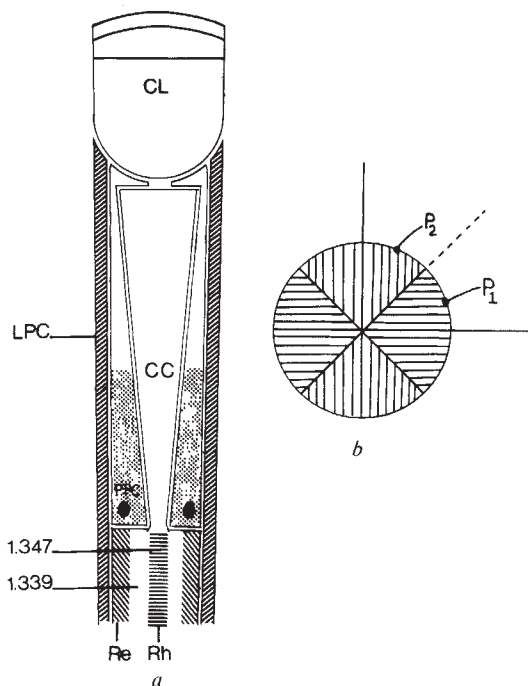


Fig. 2 a, Schematic representation of the light receptive portion of a bee ommatidium redrawn from Varela and Wiitanen⁵. CL, Cuticular lens; CC, crystalline cone; PPC, principal pigment cell; LPC, long pigment cells; Re, reticular cells; Rh, rhabdom. The numbers in the figure refer to the index of refraction measurements of Varela and Wiitanen⁵. b, A schematic representation of the worker bee rhabdom in cross section. The microvilli are shown as parallel dark lines. Two of the eight reticular cells are marked with their outputs labelled as P_1 and P_2 .

patterns known as modes which convey information about polarization and the direction of the incoming light^{5,6} (Fig. 3). The arrows represent the polarization of the electric vector. The low energy (unshaded) line of the type 2 mode is perpendicular to the plane determined by the direction of incident light and the receptor axis as shown in Fig. 3b.

To study the bee's visual sensation, we determined the amount of light absorbed by a reticular cell. Each mode has a characteristic absorption. In Fig. 4 we show that the angle of light incidence determines which of the modes of Fig. 3 predominates. A complete mathematical analysis for the light absorbed by the reticular cell is given elsewhere⁷. Qualitatively, the relative amount of absorption can be determined by superimposing on Fig. 2b the excited modes of Fig. 4. If the electric field vectors of a mode are parallel to the microvilli in

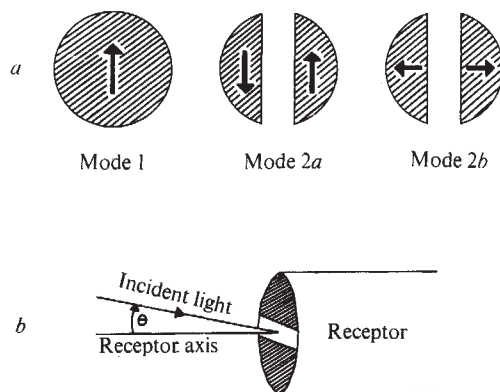


Fig. 3 a, Electromagnetic modes observed by Varela and Wiitanen⁵ on the worker bee rhabdom. Arrows represent the polarization of the electric vectors. Unshaded regions represent areas of low light intensity. The null line of mode 2 is perpendicular to the plane determined by the direction of incident light and the receptor axis as shown in b.

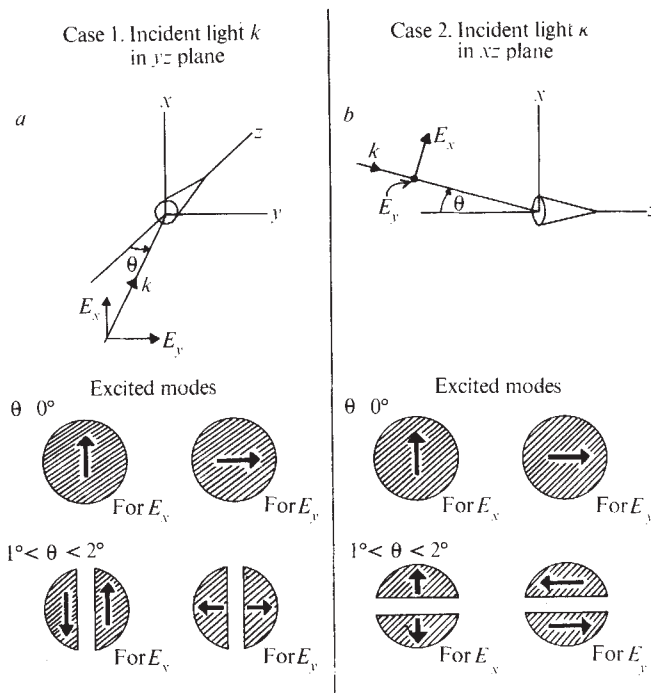


Fig. 4 Mode excitation on rhabdoms due to light obliquely incident on bee ommatidium. The incident light is assumed to be either polarized in the xz -plane (E_x) or in the yz -plane (E_y). (Only small θ values are used, $E_y(E_x)$ in case-1 (2) is nearly in the y -direction (x -direction). In the text we actually use x - and y -directions for these polarizations.) For $\theta < 1^\circ$ only the type one (Fig. 3) modes are excited. For $1^\circ < \theta < 2^\circ$ only the type two modes (Fig. 3) are efficiently excited. Note that when the electric vector is perpendicular to the plane formed by the direction of light k and the receptor axis that mode 2a of Fig. 2 is excited. As discussed in the text this mode is poorly detected. The results are taken from our electromagnetic analysis^{6,8} for ultraviolet wavelengths ($\lambda \approx 300$ nm).

a reticular cell, we call the response unity, that is $P \approx 1$. If the electric vectors of a mode are perpendicular to the microvilli, or if the unshaded (low light intensity) portion of the mode falls on the reticular cell then we have $P \ll 1$. For example, a mode like 2a of Fig. 3 is poorly detected, no matter how the rhabdom is oriented.

The relative amount of light absorbed in reticular cells 1 and 2 is presented in Table 1. We see for $\theta = 0$ that reticular cells 1 and 2 can detect polarization: for one polarization $P_1 \approx 1$, $P_2 \ll 1$ and for the other polarization $P_2 \approx 1$, $P_1 \ll 1$. This is the accepted polarization detection mechanism for the bee and is due to the alignment of microvilli. For oblique incidence, we have what seems at least at first, to be a rather peculiar effect. For the situation of Fig. 4a, x -directed polarized light is poorly detected in all reticular cells and for the situation of Fig. 4b, y -directed polarized light is poorly detected. This is because the bee rhabdom rejects mode 2a of Fig. 3 as mentioned above.

Fig. 5 illustrates the angular sensitivity of the ommatidium to off-axis light at ultraviolet wavelengths, determined

Table 1 Relative Light Absorption in Reticular Cell 1 (P_1) and in Reticular Cell 2 (P_2) of Fig. 2b caused by the Modes of Fig. 4

Mode excitation as in Fig. 4a	Mode excitation as in Fig. 4b
(A) x -direction polarized light	(A) x -direction polarized light
(A.1) $\theta = 0$; $P_1 \ll 1$, $P_2 \approx 1$	(A.1) $\theta = 0$; $P_1 \ll 1$, $P_2 \approx 1$
(A.2) $1^\circ < \theta < 2^\circ$; $P_1 \ll 1$, $P_2 \ll 1$	(A.2) $1^\circ < \theta < 2^\circ$; $P_1 \ll 1$, $P_2 \approx 1$
(B) y -direction polarized light	(B) y -direction polarized light
(B.1) $\theta = 0$; $P_1 \approx 1$, $P_2 \ll 1$	(B.1) $\theta = 0$; $P_1 \approx 1$, $P_2 \ll 1$
(B.2) $1^\circ < \theta < 2^\circ$; $P_1 \approx 1$, $P_2 \ll 1$	(B.2) $1^\circ < \theta < 2^\circ$; $P_1 \ll 1$, $P_2 \ll 1$

theoretically⁹. The null in the curve at $\theta \approx 1.5^\circ$ is due to the poor detection of mode 2a of Figs. 3 and 4a. To determine whether it would be useful for the bee to differentiate between polarized light incident on the rhabdom as in Fig. 4a as opposed to Fig. 4b, we considered the polarization properties of the sky.

Sky light is maximally polarized at an angle 90° from the Sun, with the electric vector E perpendicular to the plane defined by the following two lines: line 1 from Sun, S , to scattering region, A ; line 2 from A to observer, O (Fig. 6). This result is explained by Lipson¹⁰; detailed discussions and measurements are given by Sekera¹¹. An instrument for "fixing" on the polarization maximum should be capable of fine adjustment, at least for changes in the horizontal plane. With this in mind we now interpret the polarization effect found in the ommatidia.

In Fig. 6, ommatidium a is directed at the polarization maximum (A) and would detect the plane of polarization as described above for normal light incidence. But ommatidia b' and b receive light polarized at the interommatidial angle ψ . This corresponds to light incident with x -directed polarization as in Fig. 4a with $\theta = \psi$. If the angle $\psi \approx 1.5^\circ$ (corresponding to an absorption null (Fig. 5), only a very small portion of light from A is detected by b and b' . Then the bee would see a bright spot in its visual field corresponding to the illuminated ommatidium a surrounded (horizontally) by a dark region corresponding to ommatidia b and b' . Thus the physical properties of the rhabdom create a refined polarization compass by sharpening the directional sensitivity of ommatidia in the SOA and hence horizontal plane (Fig. 6). Neighbouring ommatidia in the plane, perpendicular to the SOA plane, strongly detect the scattered light from A and hence have no refined polarization compass properties. But this is reasonable because the bee only makes horizontal corrections in flight. It therefore seems as if the bee has a mechanism acting something like a crude two-dimensional cross hair or directional sharpening device.

We have assumed an interommatidial angle ψ of approximately 1.5° . The interommatidial angle varies over the compound eye from 0.8° to 3.57° (ref. 12). But von Frisch¹ shows that only the upper (dorsal) region of the compound eye is important for polarization navigation. In this region the average ommatidial angle is approximately 1.5° which agrees well if our electromagnetic effect is to be important. It is therefore likely that the interommatidial angle evolved to these proportions to maximize the utilization of the polarization compass.

It would be difficult to show that the bee uses the electromagnetic information as we have suggested because (a) the effect at the receptor is small because it is due to light rays off axis, and (b) even if found at the receptor level the information

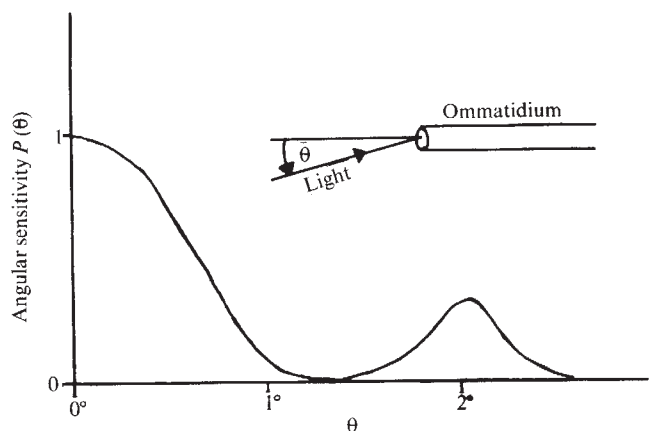


Fig. 5 Illustration of angular sensitivity of the ommatidium receiving x -directed polarized light as in Fig. 4a⁷. $P(\theta)$ can be either the response of reticular cell 1 or 2 of Fig. 2b. The null in the curve at $\theta \approx 1.5^\circ$ is due to the poor detection of mode 2a of Fig. 3 and 4a.

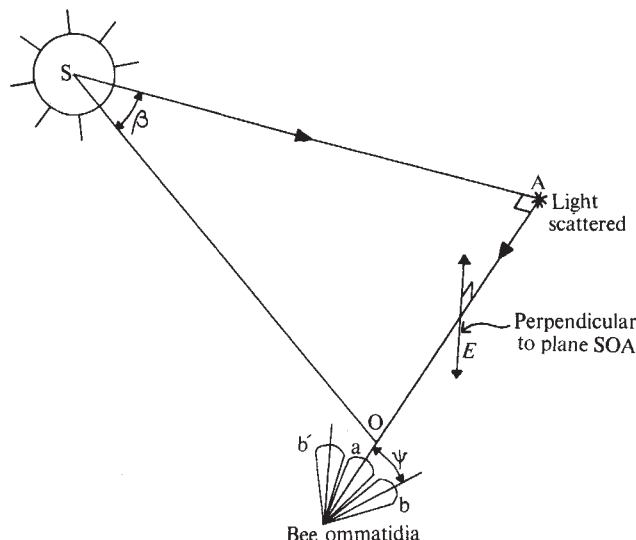


Fig. 6 Scattering of light from the Sun. *a*, *b* and *b'* represent bee ommatidia in the plane SOA. ψ is the inter-ommatidial angle. (The diagram is not drawn to scale. β is so small that SOA is also 90° . The plane SOA is inclined at some angle ϕ to the vertical, so that variations about the direction OA in that plane also represent variations in a horizontal direction. ϕ is not too large because only ommatidia in the upper or dorsal region of the compound eye detect polarized light¹.)

may not be processed by central neurones. We would expect small vibrations in the bee's flight which would cause the light from a polarization maximum to oscillate from one ommatidium to its neighbours. In this case, the bee may respond to fluctuations between maximum and minimum to enhance its detection of a "sky mark" rather than respond to the static directional fix discussed above.

It is relevant that bees respond most strongly to blue-ultraviolet electromagnetic waves^{1,15}, which are not only the most important frequencies in the blue and polarized sky light¹⁰, but are also those which produce the most pronounced electromagnetic mode rejection effect. The structure of the drone bee eye¹³ is different from that of the worker: the circular symmetry of the worker ommatidium microvilli arrangement exploited in this paper is absent in the drone, and hence the special directional effect may be absent also. The drone also has a low resistance between reticular cells¹⁴ (unlike the worker), which further reduces the mode direction. Because the drone rarely leaves the hive, there is no reason for it to have evolved special navigational aids.

Another consequence of mode propagation along the rhabdom is that an individual worker bee ommatidium is unable to detect an image¹⁶.

We thank Professor G. A. Horridge and his neurobiology group for their assistance, and Simon Laughlin for suggesting the dynamic aspects of polarization detection. C. P. is a Queen Elizabeth II research fellow.

ALLAN W. SNYDER
COLIN PASK

Research School of Physical Sciences,
Institute of Advanced Studies,
Australian National University,
PO Box 4, Canberra, ACT 2600

Received December 13, 1971; revised May 31, 1972.

- ¹ Frisch, K. V., *The Dance Language and Orientation of Bees* (Harvard University Press, Cambridge, Massachusetts, 1967).
- ² Gould, J. L., Henery, M., and MacLeod, M. C., *Science*, **169**, 544 (1970).
- ³ Lindauer, M., and Martin, H., in *Theoretical Physics and Biology*, 332 (North-Holland, Amsterdam, 1969).
- ⁴ Varela, F. G., and Porter, K. R., *J. Ultrastruct. Res.*, **29**, 236 (1969).
- ⁵ Varela, F. G., and Wiitanen, W., *J. Gen. Physiol.*, **55**, 336 (1970).
- ⁶ Snyder, A. W., and Pask, C., *Z. Verh. Physiol.* (in the press).

- ⁷ Snyder, A. W., and Pask, C., *J. Opt. Soc. Amer.* (in the press).
- ⁸ Snyder, A. W., *IEEE Trans. on Microwave Theory and Techniques*, **17**, 1138 (1969).
- ⁹ Snyder, A. W., *Z. Verh. Physiol.*, **76**, 438 (1972).
- ¹⁰ Lipson, S. G., and Lipson, H., *Optical Physics*, ch. 10 (Cambridge University Press, London, 1969).
- ¹¹ Sekera, Z., in *Handbuch der Physik*, **158**, Geophysics (Springer-Verlag, Berlin, 1957).
- ¹² Hecht, S., and Wolf, E., *J. Gen. Physiol.*, **12**, 727 (1928-29).
- ¹³ Perrelet, A., and Bauman, F., *J. Cell. Biol.*, **40**, 825 (1969).
- ¹⁴ Shaw, S. R., *Vision Res.*, **9**, 999 (1969).
- ¹⁵ Autrum, H., and von Zwehl, V., *Z. Verh. Physiol.*, **46**, 8 (1962).
- ¹⁶ Snyder, A. W., and Pask, C., *Z. Verh. Physiol.* (in the press).

Echinoids, Algal Lawn and Coral Recolonization

As a result of studies of coral reefs in the Sudanese Red Sea, I wish to suggest that the recolonization of rock areas in the rich reef crest cone^{1,2} may be facilitated by the grazing of the algal lawn by echinoids. Although the settlement of coral planulae (the free-swimming larvae) has not been investigated³, growth studies suggest that the larva is damaged by filamentous algae⁴ and therefore requires a bare rock surface for successful settlement, metamorphosis and early growth.

Even among the richest coral communities on the reef crests, up to half the surface area may be rock⁶, most of which is readily covered by filamentous algae. The size of the algal crop varies with depth reaching a maximum of 20 g m^{-2} at 3 m decreasing to 1 g m^{-2} at 10 m (see ref. 7). As this algal lawn is so ubiquitous on the reef crest, the recolonization of bare areas may depend on animals which graze on it.

On the patch reefs of the Wingate and Towartit complexes there are three dominant species of errant echinoid, but analysis has shown that all three feed on the algal lawn. Their vertical distribution is limited to the upper 6 m of reef; probably an effect of the algal crop levels. Otherwise their distribution distinctly reflects the morphological, physiological and behavioural characteristics of each species. *Echinometra mathei*⁸ is a small urchin, reaching densities of 60 per 100 m^2 , occupying the microenvironment of coral, crevices and small holes all about the reefs. *Heterocentrotus mammillatus*⁸ and *Diadema setosum*⁸ are larger echinoids found in crevices during the day-time, but, unlike *E. mathei*, emerging at night to feed on open areas of rock covered in algal lawn, totally exposed to the effects of wave action. Because of the NE prevailing wind, the NE sides of the patch reefs are exposed and the SW sides sheltered. *H. mammillatus* is found principally on the exposed side where its robust test and heavy spines give direct protection against wave action. By contrast, *D. setosum*, with fragile test and spines, is found on the sheltered sides of the reefs. Additionally, experiments indicated that *H. mammillatus* is able to adhere to a surface against an upthrust three times greater than that withstood by *D. setosum*, another factor defining their relative distributions.

Where *H. mammillatus* and *D. setosum* are present the algal lawn covers only about 20% of the rock surfaces as opposed to more complete coverage elsewhere (see Fig. 1). Grazing studies have shown that *H. mammillatus* grazes an average of 20 cm^2 per night. *H. mammillatus* and *D. setosum* are, however, only able to graze on the algal lawn if crevices for shelter during the day-time are present close to the feeding grounds.

A fourth regular echinoid, the boring urchin *Echinostrephus molaris*^{8,9} occurs frequently on the patch reefs. It bores 7.5-10 cm deep burrows, in which it is confined in adult life¹⁰. It is found in open areas of rock at maximum densities of 70 per 100 m^2 . It has been suggested that this urchin is important in contributing to the breakdown of the reef¹¹; its empty burrows are frequently occupied by *E. mathei*, and are enlarged or broken open by the presence of this urchin and by wave action, to produce crevices and holes large enough for colonization by *H. mammillatus* or *D. setosum*. These animals are

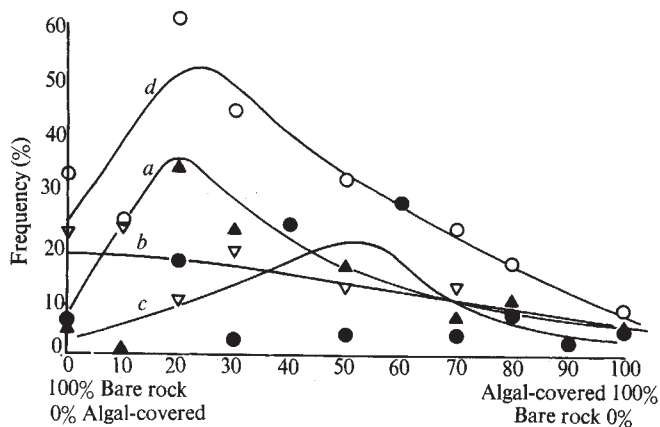
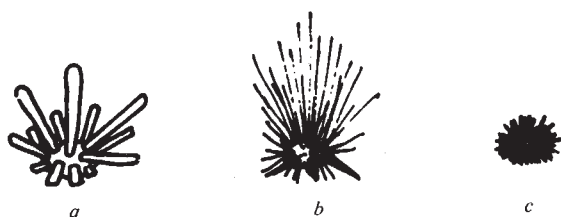


Fig. 1 Mean density of distribution as a percentage of the total sample for each species against the ratio of bare rock to algal-covered rock in the areas without coral, also expressed as a percentage. a, *Heterocentrotus mammillatus* ($\blacktriangle$); b, *Diadema setosum* (∇); c, *Echinometra mathei* ($\bullet$); d, total ($\circ$).



then able to graze the algal lawn and in this way, it is suggested, assist coral recolonization. The different stages in this succession can be seen, eventually making all areas available to the two main grazing species.

The *Echinoidea* as an order has adapted well to make maximum use of the potential algal feeding grounds on the reef. Unlike the grazing of fishes, the only other order which occurs frequently enough on the reef to affect the algal lawn, the grazing of echinoids is a very systematic scraping of the rock surface removing large areas of lawn completely and also removing superficial layers of rock^{1,2}. On these grounds it is proposed that the activity of these echinoids will have a considerable effect on colonization by any animal in which larval settlement is affected by an algal lawn. In particular, if algal growth does prevent the development of the coral larvae, the echinoids will be very important to the recolonization of rock areas by corals. In view of the present concern about coral recolonization after devastation this factor is of interest.

I thank Dr A. C. Campbell and R. F. G. Ormond for their advice and Dr C. H. Roads for facilities. This work was supported by a personal grant from Christ Church, Oxford, and by grants to the group from the Overseas Development Administration and the Royal Society.

J. K. G. DART

Trinity College,
4a Bridge Street,
Cambridge CB2 1TQ

Received April 17: revised May 24, 1972.

- ¹ Mergner, H., *Symp. Zool. Soc. London*, **28**, 141 (1971).
- ² Roads, C. H., and Ormond, R. F. G., *Rep. Third Cambridge Coral Starfish Res. Group Exped.*, 125 (1971).
- ³ Yonge, C. M., *Adv. Marine Biol.*, **1**, 209 (1963).
- ⁴ Stephenson, T. A., *Sci. Rep. Gt. Barrier Reef. Exped.*, 1928-9, **3**, 113 (1930).
- ⁵ Stephenson, A., and Stephenson, T. A., *Sci. Rep. Gt. Barrier Reef Exped.*, 1928-9, **3**, 167 (1930).
- ⁶ Barnes, J., Bellamy, D. J., Jones, D. J., and Whitton, B. A., Drew, E. A., Kenyon, L., Lythgoe, J. N., and Rosen, B. R., *Symp. Zool. Soc. London*, **28**, 87 (1971).
- ⁷ Clarke, A. M., and Rowe, F. W. E., *Monograph of Shallow-water Indo-West Pacific Echinoderms* (British Museum Natural History, London, 1970).
- ⁸ Lawrence, J. M., *Israel J. Zool.*, **19**, 175 (1970).

- ⁹ Hyman, L. H., *The Invertebrates*, **4**, 763 (McGraw-Hill, New York, 1955).
- ¹⁰ Otter, G. W., *Sci. Rep. Gt. Barrier Reef. Exped.*, 1928-9, **1**, 323 (1930).
- ¹¹ Campbell, A. C., and Ormond, R. F. G., *Symp. Zool. Soc. London*, **28**, 433 (1971).

Catastrophic Damage to Coral Reef Communities by Earthquake

THERE is interest in the disruption of coral reef communities by isolated events such as hurricanes, river floods, and torrential rainfall over reef areas^{1,2} or by seasonal fluctuations of sea level, excessive sedimentation, pollution and other human activities, and predation by *Acanthaster planci*. Dead reefs off Guadalcanal and elsewhere in the Solomon Islands have been tentatively ascribed to emersion resulting from earthquake activity³, and here I describe disruption of reef communities by earthquake near Madang, north-east New Guinea.

The New Guinea coastline between Cape Barschtsch and the Schering Peninsula is formed by a partly submerged, partly emerged barrier reef 18 km long, enclosing a lagoon 15 to 50 m deep. The mainland coast shows evidence of submergence in extensive drowned inlets, and of recent uplift in widespread raised reef terraces at +2 and +4 m (refs. 4, 5). Modern reef communities grow round islets and on shoals in the lagoon, and on the submerged barrier reef. An earthquake of October 31, 1970, force 7.1 on the Richter scale, had its epicentre at 4.9° S, 145.5° E, 50 km north-west of the reef area⁶. On the coast it was reported that the sea first receded, and then rose up to 3 m above its normal level.

I examined reef communities on the barrier reef and at fourteen sites within the lagoon between June 10 and July 2, 1971. The lagoon reefs are sheltered and are dominated by ramose and foliaceous corals, with massive corals on reef slopes. The upper surface of the barrier is dominated by more robust ramose corals, with massive corals on the seaward slope. The lagoon reefs are remarkable chiefly for the development of extensive dome-shaped colonies of tightly-packed fine-branching species, especially *Seriatopora hystrix*, *Stylophora mordax*, and *Millepora* sp.

Damage caused by the 1970 earthquake varied considerably with species and growth form. In massive corals, damage occurred either by *in situ* disintegration of the colony in species of *Porites*, *Goniopora* and *Lobophyllia* with columnar growth form, or by uprooting and downslope sliding and rolling of massive globular corals, many of which were overturned. The degree of damage to massive corals seems to depend largely on bottom slope.

Colonies of such genera as *Turbinaria*, *Echinopora*, *Echinophyllia*, *Pectinia* and *Mycodium* (vasiform foliaceous corals) were particularly fragmented, presumably because of the resistance to water movement offered by their thin folia. Such colonies generally collapsed *in situ*.

Damage to the more massive species of *Acropora*, such as *A. hyacinthus*, occurred by snapping of the holdfast and up-tilting of the colony. Such colonies remained alive except where affected by sediments. More massive species such as *A. palifera* were less affected.

Because of the sheltered situation of the lagoon reefs, delicate ramose corals are particularly widespread. In some areas earthquake effects were negligible, but in others fields of *Seriatopora*, *Stylophora*, and *Millepora* were reduced to jumbled dead branches, remaining *in situ* or showing downslope sliding. Physical breakdown was greater where massive and unattached corals rolled from above through fields of delicate ramose corals. Similar damage occurred in fields of *Acropora formosa* and other fragile *Acropora* species.

Colonies of the free-living corals *Fungia*, *Herpolitha* and *Halomitra* were scattered downslope, overturned, and in some cases lodged on top of other corals on reef slopes; they were

often inverted even on low-angle surfaces, presumably because of water agitation.

The degree of damage to coral communities varied greatly over short distances. It was greatest on reef slopes, especially in almost enclosed basins where resonance effects may have been considerable. Damage varied from less than 10% to more than 90% over 2 to 3 km, and was not obviously related to aspect. Damage to corals was primarily structural: living polyps survived even on fragmented and overturned coralla except where covered by sediment, and direct damage to polyps during the earthquake seems to have been chiefly caused by abrasion and crushing during colony movement.

In general the damage caused by the 1970 earthquake near Madang was less than that associated with major hurricanes^{7,8}. Hurricanes transport large quantities of reef debris across reef flats and down slopes, but most corals damaged by the earthquake remained *in situ* or underwent short downslope transport under gravity. Hurricane waves cause greatest damage on windward shores, but no directional trends were apparent near Madang. Debris from hurricane damage often accumulates to form supratidal bars and ridges, but no such accumulation resulted from this earthquake (Umbgrove⁹, however, recorded large reef blocks thrown on to reef flats by the 1883 Krakatau and other tsunamis). Breakdown patterns of individual corals, such as the disintegration of columnar massive corals and the fragmentation of fragile branching corals, are comparable under both earthquake and hurricane conditions and reflect similar reactions to mechanical stress; but the complete stripping of coral communities from reefs, which occurs during strong hurricanes, did not occur during this earthquake. Death of living corals as a result of mechanical damage and other forms of physiological stress, widespread during hurricanes, was of minor importance during the Madang earthquake.

A high proportion of living corals did suffer structural damage as a result of ground and water motion following the earthquake. It is possible that although immediate damage was less severe than in many hurricanes, recovery may be more delayed. Reef surfaces remain covered with irregular accumulations of interlocked dead corals, which by June 1971 were being colonized by filamentous algae. Lagoon reefs also had considerable growths of soft corals, *Sarcophyton* and *Lobophytum*, but the effects of the earthquake on these are unknown. Recolonization by corals was limited chiefly to the fresh scars left by the uprooting and movement of massive corals: small colonies a few cm in diameter of *Pocillopora*, *Stylophora*, *Seriatopora*, *Acropora* and *Millepora* were found in these situations. Because coral damage was very irregular over the reefs, it is unlikely that recovery would be delayed by the

disappearance of any species over whole sectors of the reef, as may happen during severe storms.

The effects of the Madang earthquake of 1970 were limited to superficial coral colonies, with no significant changes in the topography of reef knolls and patches, but such events must have a considerable influence on the structure and composition of reef communities. Massive corals which survive hurricanes may have an advantage over more rapidly growing but easily damaged branching corals in the hurricane belts². Earthquakes, especially in areas of high seismic activity with frequent destruction of corals, may favour rapidly growing branching species, forming colonies which are repeatedly destroyed but which rapidly regenerate. Reef fabrics in such areas probably lack the massive interlocking frameworks of robust ramose and massive corals found in more stable areas.

Most of the reef seas, especially in the Indian and central Pacific Oceans, are notably aseismic; but reefs of high species diversity are found in the East Indies, New Guinea, the Solomon Islands and New Hebrides, Fiji and Tonga, and northwards through the Philippines and Marianas, areas among the most seismically active in the world^{10,11}. Within the hurricane belts, major storms with massive destruction of corals may occur 4 or 5 times each century and affect sectors of reef 50 to 100 km wide^{7,12}. The areal extent of reef damage by large earthquakes is unknown, but in Melanesia at least their frequency is much greater than that of major hurricanes in other reef provinces. It would be useful to restudy reef communities in the East Indies, Melanesia and parts of Micronesia to determine more fully the long-term effects of catastrophic earthquakes on them.

I thank the Royal Society for support while in New Guinea, and residents of Madang for their assistance.

D. R. STODDART

Department of Geography,
Cambridge University

Received January 24, 1972.

- ¹ Stoddart, D. R., *Biol. Rev.*, **44**, 433 (1969).
- ² Stoddart, D. R., in *Applied Coastal Geomorphology* (edit. by Steers, J. A.), 155 (Cambridge Univ. Press, Cambridge, 1971).
- ³ Stoddart, D. R., *Phil. Trans. Roy. Soc. Lond.*, B, **255**, 377 (1969).
- ⁴ Panzer, W., *Z. Ges. Erdk. Berlin*, 175 (1933).
- ⁵ Krause, D. C., *Bull. Geol. Soc. Amer.*, **76**, 27 (1965).
- ⁶ Lander, J. F., *Bull. Seis. Soc. Amer.*, **61**, 483 (1971).
- ⁷ Stoddart, D. R., *Atoll Res. Bull.*, **95**, 1 (1963).
- ⁸ Blumenstock, D. I., *Atoll Res. Bull.*, **75**, 1 (1961).
- ⁹ Umbgrove, J. H. F., *Bull. Geol. Soc. Amer.*, **58**, 744 (1947).
- ¹⁰ Denham, D., *J. Geophys. Res.*, **74**, 4290 (1969).
- ¹¹ Barazangi, M., and Dorman, J., *Bull. Seis. Soc. Amer.*, **59**, 369 (1969).
- ¹² Ball, M. M., Shinn, E. A., and Stockman, K. W., *J. Geol.*, **75**, 583 (1967).

HOW TO BUY NATURE

Volumes start in January, March, May, July, September and November, but subscriptions may begin at any time.

The direct postal price per subscription is:

12 MONTHS* (52 issues per title)

	Surface mail UK and worldwide	U.S.A. and Canada
Nature (Friday)	£14	\$48
Nature + Nature Physical Science	£24	\$83
Nature + Nature New Biology	£24	\$83
All three editions	£29.50	\$108
Annual Index	£1	\$3

*Rates for shorter periods *pro rata* (minimum three months)
(Charge for delivery by air mail on application)

Editorial and Publishing Offices of NATURE

MACMILLAN JOURNALS LIMITED
4 LITTLE ESSEX STREET, LONDON WC2R 3LF
Telephone Number : 01-836 6633. Telegrams : Phusis London WC2R 3LF
Telex 262024

711 NATIONAL PRESS BUILDING
WASHINGTON DC 20004
Telephone Number : 202-737 2355. Telex 64280

Subscription Department
MACMILLAN JOURNALS LIMITED
BRUNEL ROAD, BASINGSTOKE, HANTS
Telephone Number : Basingstoke 29242

American display advertisements
NATURE SCIENTIFIC PUBLICATIONS INC
711 NATIONAL PRESS BUILDING
WASHINGTON DC 20004

All other advertisements
T. G. SCOTT & SON, LIMITED
1 CLEMENT'S INN, LONDON WC2A 2ED
Telephone : 01-242 6264/01-405 4743
Telegrams : Textualist London WC2A 2ED

Registered as a newspaper at the Post Office

Copyright © Macmillan Journals Limited, September 1 1972

BOOK REVIEWS

A Boswell's Rutherford

Rutherford: Recollections of the Cambridge Days. By Mark Oliphant. Pp. xii+158+36 photographs. (Elsevier: Amsterdam, London and New York, 1972.) Dfl.22.50; \$7.

THE Centenary of Rutherford's birth was celebrated in New Zealand, Canada, Russia and Britain last year. The Addresses delivered to the Royal Society have not yet appeared in print; they concern mostly impressions of the Cambridge period and here is another contribution by Sir Mark Oliphant, a delightful volume of reminiscences by one who was accepted by the Rutherfords almost as an adopted son. For this reason it is more personal than are most biographies and as such is an important addition to the sum total of our knowledge of one of the world's greatest physicists. Sir Henry Tizard, the Secretary of the DSIR during Rutherford's Cambridge days, remarked that it was a pity that Rutherford had no Boswell, that no one had made notes of his ways, his sayings, his faults and his virtues: Oliphant's privileged position gave him a unique opportunity to observe Rutherford's idiosyncracies and we are grateful to Oliphant for publishing very many recollections and anecdotes of the kind Boswell would have recorded.

The official biography of Rutherford was written by Professor Eve, that for the Royal Society by Professor Eve and Sir James Chadwick, Sir Ernest Marsden wrote for the Royal Society of New Zealand, and Sir Henry Tizard wrote for the DNB. The Manchester period of Rutherford's life was well recorded at the 1961 Conference to mark the Jubilee of the discovery of the atomic nucleus; some of his students of that period told many interesting stories of Rutherford but as Professor H. R. Robinson then said we cannot

lections and contributes very much new material.

Oliphant arrived in England in 1925 and on entering the Cavendish "I was received genially by a large, rather florid man, who reminded me forcibly of the keeper of the general store and post office in a little village behind Adelaide where I had spent my childhood. Rutherford made me feel welcome and at ease at once." This was surely one of his most endearing characteristics; whenever he went round the Laboratory he greeted each research student with a friendly question about progress of work phrased in such a way as to banish fear or embarrassment and thus evoked a detailed discussion in the freest terms. Just occasionally he would express displeasure, but his strictures were generally deserved: when they were not, his greatness of character was shown by his generous apology, like the one he made to Oliphant after his rage over Oliphant's acceptance of the Chair in Birmingham in 1936; "Go and be damned to you." Many examples are given of this kindly approach; in the garden of Newnham Cottage, in a ship's engine room talking with stokers and greasers, when chairing a meeting and when trying to console Kapitza who had been forcibly detained in Russia in 1934.

The age of "sealing-wax and string" had almost gone but the Laboratory was run on a very modest budget, only £16,000 from the university in 1935 including "£30 on an international conference". Rutherford had had to be very careful with money as a young man—he had had to borrow the money for his fare from New Zealand to England in 1895—and even when the Laboratory was famous he preferred to run it on a shoe-string. The Master of Caius College had persuaded some wealthy friends to offer Rutherford a private grant for research but there is no evidence that he used it and when Oliphant once said to him "it is a pity nobody has given you a gram of radium" (as had been given to Madame Curie), he replied "I am very glad nobody did; how on earth could I justify the use of a whole gram of radium." As Sir James Chadwick wrote in *Nature* shortly after Rutherford's death, "he knew his worth

but he was and remained, amidst his many honours, innately modest."

There is a full account of "The Crocodile". Kapitza's new Mond Laboratory was to have a plaque carved by Eric Gill in the entrance hall and a decorative theme carved just outside the door: this theme was of Kapitza's choosing—a crocodile. Kapitza apparently was always afraid of having his head snapped off, as by a crocodile, but an alternative explanation was that, like the crocodile in Peter Pan, Rutherford could be heard approaching long before he was visible. As recently as last September Kapitza would not tell the reviewer of this book why he had chosen this theme and the book does not answer the question. But the plaque portrayed Rutherford with a Jewish nose and it elicited much adverse criticism: the correspondence on the matter is here quoted extensively. But the person who minded least was Rutherford himself; he merely remarked that his name should be "Rutherfordmond".

The chapter which old Cavendish men will turn to first is the very personal story about "Home, Holidays, Politics": the tragedy of his daughter's death in 1930 just as he had heard privately of his forthcoming elevation to the Peerage, the sadness at the death of his mother at the age of 92, the joys of the country cottages in Wales and later in Wiltshire, Lady Rutherford's gardening and "Ern's" politics; to Oliphant he said "that (military tattoo) expressed the true spirit of England far better than all the nonsense you read in *The New Statesman*." He did not seem to have any absorbing hobby except reading, and in particular, reading history.

Sir Mark concludes "In all ways except in his science, Rutherford was an exceptionally ordinary man in both appearance and character. Quite naturally, he was friendly with all men and quarrelled with none. His success came as much from complete dedication to his work as from his innate ability, so that even the average student was inspired to emulate him. Yet there was something in him which raised him high above others and put him in the company of the greatest of men."

T. E. ALLIBONE

"parcel out
His intellect by geometric rules,
Split like a province into round and
square."

The Centenary celebrations were enlivened by stories told by students of the Cavendish Laboratory and the present volume greatly enhances those recol-

Manufacturing Food

Technological Eating or Where Does the Fish Finger Point. By Magnus Pyke. Pp. 114+21 plates. (John Murray: London, February 1972.) £2.50.

MODERN food technology is a recent development but it has grown and made its impact with amazing speed. We are so familiar with dried, pre-packaged, frozen and instant foods, supermarkets and vending machines that we forget that it is not many years since most housewives baked their own bread, the breakfast cereal was unknown and no one had a refrigerator.

Some of these developments have been and are being directed by social circumstances—the disappearance of domestic help, the working housewife, increasing affluence, the skills of the food producer and the newer techniques of persuasion—at the same time developments in food technology are themselves creating major changes. Technologies of all kinds affect human behaviour and Dr Pyke points out that the effects of food technology are no less than those of communication satellites and nuclear power stations. The slave trade, the history of South America and social structure of many parts of West Africa and the Pacific were greatly affected by food technology.

The slave trade, continuing for three centuries, was created as the result of the need to provide workers for the sugar industry. The eventual loss of the slaves led to the development of the sugar beet in the temperate zone, with long-term effects on the livelihood of cane-sugar producing countries. Bananas took their place as a major food commodity in the late 19th and early 20th centuries but only after the major exploitation of the virgin lands of half a dozen South American States. Margarine, a starting point in modern food technology, led to a search for vegetable oils and changed the complex social life of vast areas of West Africa and the Pacific.

On a more personal scale developments in food technology have altered many of our social habits and cultural traditions, which were often based on a link between man and the soil. When foods come from the factory rather than the farm they tend to be regarded as industrial products rather than cultural factors. When the avoidance of milk produced by cows of another family, as among some Bantu communities, is overcome by detribalized spray-dried milk, the result may well be an improvement in nutritional status. When the avoidance of pork can be overcome by eating simulated pork made from textured vegetable protein the morality is at least open to debate.

Another result, inevitable in the modern business world, is the gross enlargement of the food business. Dr Pyke selects an example of agribusiness near his home, from among the raspberry growers near Dundee. There, the largest acreage of raspberries in the world, the fruit was grown by farmers, collected by school children and sold to jam makers. Now the farmers have gone into business—the plant breeders are developing varieties suitable for mechanical harvesting, the engineers are designing the equipment and the farmers are going into jam-making themselves. The chicken farmer, the supermarkets that manufacture and even grow their own supplies, the mergers and the vertical integration, give rise to the multinational firms that we know today, no less in the food industry than in the petrol and motor car industries. This has changed our regard for food manufacture, much to the regret of many who surround food production with the aura of nature, countryside, leisure and peace, while food processing is regarded as a nasty business. As Dr Pyke says, food manufacture, in many respects, is no different from the manufacture of any other commodity. A side issue is the growing uniformity of food on the international scale—the same breakfast cereal is eaten around the western world and is becoming popular in the east, the same wrapped and branded commodities sold in our supermarkets find their way into the oriental bazaars, and the same baby foods, tea and biscuits, not to mention Cola drinks and whisky.

When we can spin soy beans to resemble meat, when the entire world's needs of protein can be produced in a volume of 10^{13} litres, when the food needs of an entire family for a year can be stored outside the kitchen window in a deep freeze, Dr Pyke asks "where does the fish-finger point?" Will the growth continue—600 million fish-fingers, 1.5 million tons of frozen vegetables, milk that will store for months? He accepts the fallacy of extrapolation and the unpredictability of events; certainly the speed of change makes any forecasts vulnerable. He suggests that growing affluence allows people to choose how they wish to behave and that this may well reverse the current trend to uniformity. Do-it-yourself is not always the result of shortage of money and applies to food preparation as well as to paperhanging. With the same illogic with which the Englishman will eat cow but not dog, will hang game but not chicken, we may well, as a special celebration, forswear the frozen dinner and go to the trouble and loving care of roasting a shoulder of lamb—from meat taken from the deep freeze.

ARNOLD E. BENDER

Electron Deficiency

Electron Deficient Compounds. By K. Wade. Pp. vi+203. (Thomas Nelson: London, November 1971.) £3.95 boards, £1.95 paper.

THERE is an increasing trend towards the publication of shorter specialized student texts, but for such texts to be wholeheartedly recommended they must be unique in the area they cover, and exemplary in presentation. With only minor reservations this book fulfils both these requirements admirably. The discussion of electron deficient compounds appears at some stage in the vast majority of undergraduate courses, and this publication is timely in that it presents, for the first time, a comprehensive coverage of the area at this level.

After a short, but useful, introductory chapter which emphasizes the type of compounds to which the term "electron deficient" can be sensibly applied, there follows six chapters. The first three of these deal with the boron hydrides and their derivatives, covering the structure and bonding of diborane, for the higher boranes and polyhedral systems, and then a general discussion of the reactions of the boron hydrides. The last three chapters deal successively with associated metal hydrides, alkyls and aryls, the carboranes, and finally a pertinent section on other examples of electron deficient compounds including transition metal clusters.

The balance of each chapter is well maintained with the possible exception of chapter 3 where, in determining the charge distribution in the higher boron hydrides, undue emphasis is placed on the simple electron counting method within the various types of three-centre electron-pair bonds. The techniques used in the investigation of such compounds are adequately covered although the use of mass spectrometry is not discussed. Two useful features are the list of references for further reading, and the series of problems, given at the end of each chapter. The answers to the problems are outlined at the end of the book.

The diagrams are clear and well presented, and there are very few printing errors. The ^{11}B nuclear magnetic resonance spectrum of decaborane on page 40 could be replaced with advantage by a more recent spectrum, and some indication of the frequency at which the spectra are recorded would be useful.

These few trivial qualifications apart, however, this is a reasonably priced book which students at both the undergraduate and research level will find extremely valuable, and it can be highly recommended as an authoritative introductory text for the study of the title compounds. M. G. H. WALLBRIDGE

Filtering Membranes

Permselective Membranes. Edited by C. E. Rodgers. (Selected papers presented at the American Chemical Society Symposium, New York, September 1969.) Pp. xi+206. (Marcel Dekker: New York, September 1971.) \$14.50. (First published in *Journal of Macromolecular Science*, Vol. 5B, No. 1, 1971.)

I READ this collection of diverse contributions to various aspects of membrane transport mechanisms hoping that there might be found within these pages well-defined achievements or clear statements of problems seeking solution.

I found neither. The publication contains random contributions on mechanistic rate studies of gas transfer in acrylates, block copolymers, silicone rubbers and irradiated fluorocarbon films. It also contains yet more fundamental studies of solute and water transport in hyper-filtration membranes, and anyone however vaguely acquainted with the field must know by now that $P=D\alpha S$.

The editor states in the preface of this volume that "An ultimate aim of research in this field is to establish mechanisms and laws relating solubility and transport in multi-component systems with the molecular properties and characteristics of the components". True and worthy, no doubt, but for the permselective membrane field to achieve true credibility in areas other than the artificial kidney (scantly mentioned), another and perhaps more pressing and immediate aim of research into membrane permeability would be to tackle immediate problems where permselective membranes could make immediate contributions.

In the desalination field, cellulose acetate lacks certain mechanical properties to withstand creep when solutions are pumped over its surface at fluid velocities which determine possible economic operation of desalting plant. It therefore lacks credibility. But cellulose acetate membranes could perhaps replace sand and gravel filters for sewage and organic component removal if open-pore membranes could produce fluxes approaching the 120 (UK) gallons square foot per hour given by gravel filters. An immediate application for a permselective membrane? We need to know.

In the medical field, is it possible fully to oxygenate 5 l. of blood per minute, simultaneously exchanging carbon dioxide to acceptable partial pressures with a permselective silicone membrane in view of the fact that direct process equipment produces extremely serious blood damage? We need to know. These questions remain unanswered by this book, in spite of the

fundamental knowledge which has been gathered in the past decade on membrane systems for separation processes.

W. M. MUIR

Free Radicals in vivo

Free Radical Mechanisms in Tissue Injury. By T. F. Slater. Pp. 283. (Pion: London, 1972.) £4.80. Students' edition, £3.

IN these days of ever increasing specialization in science any attempt to bridge the gaps between traditional disciplines is welcome. An example of bringing together some of the interests of physical chemistry and physiology is provided by this monograph. Although this is the expressed aim of the author, he has not used the most effective way to achieve his purpose. As he himself points out, it is difficult to gain the confidence of any specialist by discussing his subject incompletely. It is obviously a near impossible task to describe in some eighty pages free radical reactions including techniques. Inevitably the author's selection is unlikely to appeal to everybody. Inevitably and perhaps wrongly, the reader with some knowledge in free radical work suspects that the treatment of the physiological points may suffer from similar defects. An even slightly discouraged reader is unlikely to become interested and stimulated enough to want to go into the new field. In my view the author would have been more successful if he had provided two shorter monographs in which the ration of one discipline was confined to less than 10% of that of the other. A ten or twenty page introduction to the specialty of the reader with reference to standard textbooks would be less likely to produce an unwanted reaction.

In the introduction of two to three pages the author provides a definition of free radicals and, surprisingly, a description of membranes and their importance to living tissues. Their relevance to this monograph is justified as follows: since membranes often contain unsaturated lipids and are surrounded by oxygen rich, metal containing fluids they are susceptible to peroxidative attacks which involve free radical reactions. After these short remarks the reader is given a description of free radicals, their production, chemical reactions, methods of detection, free radical scavengers, their occurrence in biological materials and some of their deleterious reactions with biologically important materials. This part, about a third of the monograph, is lucid and cohesive. The author then

builds up in a few pages to his real interest, which is the action of carbon tetrachloride on the liver, leading to fatty degeneration and necrosis. Hepatotoxic effects of alcohol follow with a predominantly physiological and biochemical discussion giving very little information on free radical involvement. The liver and its reactions take up more than a third of the book. Free radical damage to biological membranes is discussed in some 20 pages which no doubt should appeal to free radical chemists, providing a concise and interesting introduction to this field. Free radical involvement in chemical carcinogenesis and radiation-induced tissue damage get only some thirty-five pages and are correspondingly sketchy although this subject fits the title of the book very well. The book ends with a few pages on iron and copper toxicity which unfortunately are only loosely linked to free radical mechanisms.

The style of writing is precise and the literature citations are adequate for further information in unfamiliar fields. The book was printed by a photolitho process but is well produced with clear illustrations and a good subject index. For the non-initiated but interested reader it would have helped to get a more obviously cohesive choice of examples from the biological field with perhaps more indications of possible physico-chemical approaches to the problems involved. With some courage, however, the author has produced a useful and readable selection of chemical and physiological facts. Those scientists already working in this interdisciplinary area will appreciate it greatly. How many new converts this monograph is going to make only the future will tell. But even if only a handful of good physical chemists are encouraged to enter the physiological field they should be in a position to make a contribution. Let us prevent science from drifting apart and get some help from books like this to follow the literature in areas outside our own specialties.

M. EBERT

Archaeology in Context

Environment and Archaeology: An Ecological Approach to Pre-history. By Karl W. Butzer. Second edition. Pp. xxvi+703. (Methuen: London, May 1972.) £8.

WHEN Professor Butzer originally wrote this book he set himself a tremendous task: a review of the wide range of techniques used in studying the Ice Age and a series of selected examples of geographical reconstructions of certain areas and certain phases in the Pleisto-

cene. The book became indispensable to a wide range of students interested in the Pleistocene, so that a second and expanded edition is welcome.

When one compares the two editions the first half on techniques turns out to have increased from 265 only to about 270 pages; but the second half (comprising Parts V and VI, the examples) has been extensively altered and rewritten: it has increased from just over 200 pages to just under 350. The new edition is not as well produced as the old one. The reproduction of some line diagrams is distinctly dirty, and has deteriorated between editions as can be seen from a comparison of the forty-nine diagrams included in the first half of the book. The heaviness of the type is somewhat variable even on a single page. In fact the original edition looks a better book.

Yet one is again filled with admiration for the tremendous range of Professor Butzer's knowledge. The book is heavily documented, seventy-five pages of references for 600 pages of text. The field of knowledge covered is full of uncertainties and personal hunches and at times the compressed, heavily documented style gives a much greater air of certainty than is in fact justified. Yet in other places the author's scepticism is profound; for example, the beginning of chapter 20. I myself would have preferred a somewhat more leisurely treatment of a smaller field (possibly fewer examples) to give a non-expert more discussion of the value of the various types of evidence. This would probably have meant a more tentative and descriptive and perhaps less erudite approach that might well not have been what the author had in mind.

I am happy to report that although the "geography" of the original subtitle has now become "ecology" this merely reflects a fashion in words and not a radical change of subject.

B. W. SPARKS

African Butterflies

Life Histories of South African Lycaenid Butterflies. By G. C. Clark and C. G. C. Dickson. Pp. xvi+272+108 plates. (Purnell: Cape Town, Johannesburg, London and New York, May 1972.) £9.

SPECIFIC identification of adults in a group of more than 148,000 species presents frequent difficulties. Identification of the immature stages of Lepidoptera is at present impossible for most species: relatively few have been reared, illustrated or preserved in collections. This handsome book goes a small part

of the way towards filling this large gap in our knowledge.

The authors illustrate, mostly for the first time, the early stages of 125 of the 200 or so South African Lycaenidae. The eggs and each larval instar are figured, together with the pupa and adult butterfly. The colour-plates, which are based on the superb paintings of Clark, are generally good, although a harder paper would have improved the definition. The numbers accompanying each group of figures correspond to the species-number in the text, but are not invariably consecutive: for example, 106 and 107 appear after the plate of species Nos. 105 and 108. Concise information is given about larval feeding habits, external characters of the eggs, larvae and pupae, parasites of the early stages, flight period and behaviour of the adult, and species distribution in South Africa and (briefly) elsewhere. Many of these data are the result of extensive fieldwork by the authors, who were the first to trace the complicated life-histories of several species, a number of which are attended by ants or spend the whole or part of their larval existence in ants' nests where they are probably predaceous on ant larvae. One species has been reared on Psyllidae (Homoptera); others feed on Coccidae (Homoptera) during the early instars.

A key to the larvae would have made a useful addition to this book—even if only to genera—as would a note in the introduction drawing the attention of readers (abstracters in particular) to the description of a new species on page 7, and to the several changes in nomenclature, many in disagreement with Stempffer (*Bull. Br. Mus. Nat. Hist. (Ent.)*, Suppl. 10; 1967): for example, *Argiolaus* is elevated to generic status and *zeuxo* is transferred to *Poecilmitis*.

The authors conclude their introduction with the hope that others will be stimulated to continue their investigations into the life-histories of Lycaenidae: their book is not only stimulating in this respect but indispensable. The ecology and physiology of Lycaenid larva and ant association offer fascinating fields of study. Possible selective advantages enjoyed by ant-nest inhabiting larvae, perhaps by reduction from predator attack or protection from fire, could be worth investigation.

A. WATSON

Sedimentology

Carbonate Sediments and Their Diagenesis. By R. G. C. Bathurst. Pp. xix+620. (Elsevier: Amsterdam, London and New York, 1971.) Dfl. 90; \$25.75.

THIS book is concerned with modern carbonate sediments and the processes which they undergo following deposition. The first half of the book deals with carbonate particles that form in modern environments, especially skeletal fragments derived from the disintegration of the shells of organisms, and ooids, pisolites, peloids, and aragonite needles which compose lime-mud. The environments in which these particles form, mostly tropical shallow seas, including the Bahamas, the shelf of southern Florida, the Gulf of Batabana (Cuba), the Persian Gulf, and the shelf of British Honduras, serve as models to illustrate the interaction of the many parameters that are responsible for the formation of carbonate particles.

The second half of the book concerns itself with the subject of diagenesis, a term which applies to post-depositional alteration which particles undergo. This subject relates by and large to processes by which particles become cemented together to form rock. Cementation or "lithification" is one of the most important natural processes occurring at or near the surface of the Earth. Hence how carbonate particles cement together to form carbonate rock or limestone is an important concern of this book. The process of cementation also has economic implications. Porous limestones serve as reservoirs for oil, gas, and water; thus the degree of cementation and the reasons for elimination of pore space or the creation of secondary pores become important as principles of economic geology.

Although this book has been devoted to Bruno Sander of Austria, an inspiring pioneer in the study of carbonate fabric, much of the material presented is based on the work of geologists working in major oil companies. During the 1950s and early 1960s research in carbonate sediments advanced rapidly under the economic incentive of the needs to explore for oil and gas in limestones. Hence the major oil companies formed research groups in carbonate sedimentology and finally made available to the profession through the medium of publication the results of this fundamental research.

The book consists of thirteen chapters, a glossary of sedimentological terms not defined in the text, a reference section, and the index. It is an outstanding and stimulating synthesis of the field of carbonate sedimentology. It has been written for the graduate student. After digesting the contents of this book a student will be well versed and up to date in this field, even if he has no prior knowledge of carbonate sedimentology. The illustrations are excellent. It is the most outstanding volume of Elsevier's series "Developments in Sedimentology".

GERALD M. FRIEDMAN